

Information Retrieval-Mehrwertdienste für Digitale Bibliotheken: Crosskonkordanzen und Bradfordizing

Mayr, Philipp

Veröffentlichungsversion / Published Version
Dissertation / phd thesis

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Mayr, P. (2010). *Information Retrieval-Mehrwertdienste für Digitale Bibliotheken: Crosskonkordanzen und Bradfordizing*. (GESIS-Schriftenreihe, 5). Bonn: GESIS - Leibniz-Institut für Sozialwissenschaften. <https://doi.org/10.21241/ssoar.26122>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see:
<https://creativecommons.org/licenses/by-nc/4.0>

Information Retrieval- Mehrwertdienste für Digitale Bibliotheken

Crosskonkordanzen und Bradfordizing

Philipp Mayr

Information Retrieval-Mehrwertdienste für Digitale Bibliotheken

GESIS-Schriftenreihe

herausgegeben von GESIS – Leibniz-Institut für Sozialwissenschaften

Band 5

Philipp Mayr

Information Retrieval-Mehrwertdienste für Digitale Bibliotheken

Crosskonkordanzen und Bradfordizing

Die vorliegende Arbeit wurde von der Philosophischen Fakultät I der Humboldt-Universität zu Berlin im Jahr 2009 als Dissertation zur Erlangung des akademischen Grades Doctor philosophiae (Dr. phil.) angenommen.

Philipp Mayr

**Information Retrieval-Mehrwertdienste
für Digitale Bibliotheken
Crosskonkordanzen und Bradfordizing**

Bibliographische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN 978-3-86819-012-0

ISSN 1869-2869

Herausgeber,

Druck u. Vertrieb: GESIS – Leibniz-Institut für Sozialwissenschaften
Lennéstraße 30, 53113 Bonn, Tel.: 0228 / 22 81 -0
info@gesis.org

Printed in Germany

©2010 GESIS – Leibniz-Institut für Sozialwissenschaften, Bonn. Alle Rechte vorbehalten. Insbesondere ist die Überführung in maschinenlesbare Form sowie das Speichern in Informationssystemen, auch auszugsweise, nur mit schriftlicher Einwilligung von GESIS gestattet.

Inhalt

| | |
|---|----|
| Zusammenfassung | 9 |
| Abstract | 10 |
| Widmung | 11 |
| Danksagung | 13 |
| 1 Einleitung | 15 |
| 1.1 Thematische Einordnung | 16 |
| 1.2 Motivation und Ziele | 18 |
| 1.3 Forschungsfragen | 19 |
| 1.4 Aufbau der Arbeit | 20 |
| 2 Bibliographische Datenbanken, Portale und Suchmaschinen | 23 |
| 2.1 Ausgangssituation | 25 |
| 2.2 Allgemeinsuche versus Fachzugang | 27 |
| 2.2.1 Analytische LiteraturoAuswertung | 28 |
| 2.2.2 Fazit | 34 |
| 2.3 Informationstypen | 34 |
| 2.3.1 Literatur | 34 |
| 2.3.2 Internetquellen | 35 |
| 2.3.3 Primärinformation | 36 |
| 2.3.4 Fazit | 37 |
| 2.4 Informationsangebote | 37 |
| 2.4.1 Fachdatenbanken & Bibliographien | 38 |
| 2.4.2 OPACs | 38 |
| 2.4.3 Sammlung von Internetquellen | 39 |
| 2.4.4 Dokumentenserver | 40 |
| 2.4.5 Digitalisate | 40 |
| 2.5 Informationsanbieter | 41 |
| 2.5.1 Bibliotheken | 41 |
| 2.5.2 Fachinformationsanbieter | 42 |
| 2.5.3 Kommerzielle Anbieter | 42 |
| 2.5.4 Fachinformation im Internet | 42 |
| 2.6 Elektronische Publikationen | 43 |
| 2.6.1 Open Access & Self-archiving | 45 |
| 2.6.2 Integration von elektronischen Publikationen | 46 |

| | | |
|---------|--|-----|
| 2.7 | Google Scholar als akademische Suchmaschine | 47 |
| 3 | Heterogenität und Information Retrieval | 55 |
| 3.1 | Schalenmodell | 55 |
| 3.2 | Projekt KoMoHe | 58 |
| 3.2.1 | Heterogenitätsbehandlung | 60 |
| 3.2.2 | Erstellung von Crosskonkordanzen | 64 |
| 3.2.2.1 | Einsatz von Crosskonkordanzen am Beispiel infoconnex | 66 |
| 3.2.2.2 | Indirekte Term-Transformation | 68 |
| 3.2.3 | Ergebnisse der Mapping-Initiative | 70 |
| 3.2.4 | Implementation der Crosskonkordanzen | 77 |
| 3.2.5 | Heterogenitätsservice (HTS) | 79 |
| 4 | Bradford's Law of Scattering | 85 |
| 4.1 | Einführung des BLS | 86 |
| 4.2 | Bradfords Daten | 88 |
| 4.3 | Einsatzmöglichkeiten | 92 |
| 4.4 | Bradfordizing-Effekt | 94 |
| 5 | Material und Methode | 97 |
| 5.1 | Bradfordizing | 98 |
| 5.2 | Evaluation | 101 |
| 5.2.1 | Evaluation der Term-Transformationen | 102 |
| 5.2.2 | Evaluation des Re-Rankings | 109 |
| 5.3 | Testkorpora | 111 |
| 5.3.1 | GIRT-Korpus | 111 |
| 5.3.1 | KoMoHe-Kollektionen | 112 |
| 5.4 | Topics | 113 |
| 5.4.1 | CLEF-Topics | 114 |
| 5.4.2 | KoMoHe-Topics | 120 |
| 5.5 | Pooling der Dokumente | 123 |
| 5.6 | Relevanz-Bewertung | 123 |
| 5.7 | Evaluationsmaße | 126 |
| 5.7.1 | Precision- und Recall-Berechnung | 127 |
| 5.7.2 | Beispielsberechnung | 128 |
| 5.8 | Signifikanztest | 129 |
| 5.9 | Qualitatives Kurzinterview | 130 |

| | | |
|-------|---|-----|
| 6 | Effekte von Crosskonkordanzen und Bradfordizing | 133 |
| 6.1 | Überlappung und Termzugewinn durch Crosskonkordanzen („soFid-Experiment“) | 133 |
| 6.2 | Bradfordizing bei Ein-Wort-Anfragen | 138 |
| 6.3 | Bradfordizing bei Deskriptoren-Kombinationen | 140 |
| 6.4 | Analyse der „Informetrics“-Literatur in der LISA-Datenbank | 142 |
| 6.5 | Periodizität der Zeitschriften | 146 |
| 6.6 | Informetrische Analyse von Top Level Domains | 148 |
| 7 | Ergebnisse | 153 |
| 7.1 | Effekte der Crosskonkordanzen beim IR | 153 |
| 7.1.1 | Test A: Suche mit kontrollierten Termen | 154 |
| 7.1.2 | Test B: Freitext-Suche | 156 |
| 7.1.3 | Zusammenfassung | 157 |
| 7.2 | Bradford-Verteilungen in Fachdatenbanken | 158 |
| 7.2.1 | Verteilung der Zeitschriftenartikel auf Bradford-Zonen | 159 |
| 7.2.2 | Verteilung der Monographien auf Bradford-Zonen | 165 |
| 7.2.3 | Zusammenfassung | 169 |
| 7.3 | Befragung zu Bradfordizing als Mehrwertdienst | 170 |
| 7.3.1 | Interview zu Kernzeitschriften | 170 |
| 7.3.2 | Zusammenfassung | 173 |
| 7.4 | Effekte von Bradfordizing beim IR | 174 |
| 7.4.1 | Precision innerhalb der Bradford-Zonen | 174 |
| 7.4.2 | Zeitschriftenartikel | 174 |
| 7.4.3 | Monographien | 181 |
| 7.4.4 | Zusammenfassung | 186 |
| 8 | Diskussion | 187 |
| 8.1 | Effekte von Crosskonkordanzen für das IR | 187 |
| 8.2 | Bradford-Verteilungen für Zeitschriften- und Monographienliteratur. | 193 |
| 8.3 | Mehrwert durch Bradfordizing | 196 |
| 8.4 | Relevanzvorteil durch Bradfordizing bei Zeitschriftenartikeln und Monographien | 198 |
| 9 | Ausblick | 207 |
| 10 | Thesen | 211 |
| | Literaturverzeichnis | 213 |

Anhang

| | |
|--|-----|
| Beispielfragebogen. | 233 |
| Signifikanztests | 236 |
| Signifikanztest CLEF 2003-2007 (Zeitschriftenartikel) | 236 |
| Signifikanztest CLEF 2005-2007 (Monographien). | 238 |
| Signifikanztest CLEF 2003 (Zeitschriftenartikel) | 240 |
| Signifikanztest CLEF 2004 (Zeitschriftenartikel) | 242 |
| Signifikanztest CLEF 2005 (Zeitschriftenartikel) | 244 |
| Signifikanztest CLEF 2006 (Zeitschriftenartikel) | 246 |
| Signifikanztest CLEF 2007 (Zeitschriftenartikel) | 248 |
| Signifikanztest CLEF 2005 (Monographien) | 250 |
| Signifikanztest CLEF 2006 (Monographien) | 252 |
| Signifikanztest CLEF 2007 (Monographien) | 254 |
| Signifikanztest KoMoHe (Zeitschriftenartikel). | 256 |
| Signifikanztest KoMoHe (Monographien) | 258 |
| Signifikanztest KoMoHe IR-Test | 260 |
| Kontrollierte Vokabulare | 262 |
| Ausschnitte aus der Crosskonkordanz Thesaurus Sozialwissenschaften (TheSoz) zu Standard Thesaurus Wirtschaft (STW). | 264 |
| Ausschnitte aus der Crosskonkordanz Psynindex Terms zu Medical Subject Headings (MeSH) | 265 |
| Indirekte Term-Transformationen | 266 |

Zusammenfassung

Trotz großer Dokumentmengen für datenbankübergreifende Literaturrecherchen erwarten akademische Nutzer einen möglichst hohen Anteil an relevanten und qualitativen Dokumenten in den Trefferergebnissen. Insbesondere die Reihenfolge und Struktur der gelisteten Ergebnisse (Ranking) spielt, neben dem direkten Volltextzugriff auf die Dokumente, inzwischen eine entscheidende Rolle beim Design von Suchsystemen. Nutzer erwarten weiterhin flexible Informationssysteme, die es unter anderem zulassen, Einfluss auf das Ranking der Dokumente zu nehmen bzw. alternative Rankingverfahren zu verwenden.

In dieser Arbeit werden zwei Mehrwertverfahren für Suchsysteme vorgestellt, die die typischen Probleme bei der Recherche nach wissenschaftlicher Literatur behandeln und damit die Recheresituation messbar verbessern können. Die beiden Mehrwertdienste semantische Heterogenitätsbehandlung am Beispiel Crosskonkordanzen und Re-Ranking auf Basis von Bradfordizing, die in unterschiedlichen Phasen der Suche zum Einsatz kommen, werden hier ausführlich beschrieben und im empirischen Teil der Arbeit bzgl. der Effektivität für typische fachbezogene Recherchen evaluiert.

Vorrangiges Ziel der Promotion ist es, zu untersuchen, ob das hier vorgestellte alternative Re-Rankingverfahren Bradfordizing im Anwendungsbereich bibliographischer Datenbanken zum einen operabel ist und zum anderen voraussichtlich gewinnbringend in Informationssystemen eingesetzt und dem Nutzer angeboten werden kann.

Für die Tests wurden Fragestellungen und Daten aus zwei Evaluationsprojekten (CLEF und KoMoHe) verwendet. Die intellektuell bewerteten Dokumente stammen aus insgesamt sieben wissenschaftlichen Fachdatenbanken der Fächer Sozialwissenschaften, Politikwissenschaft, Wirtschaftswissenschaften, Psychologie und Medizin.

Die Evaluation der Crosskonkordanzen (insgesamt 82 Fragestellungen) zeigt, dass sich die Retrievalergebnisse signifikant für alle Crosskonkordanzen verbessern; es zeigt sich zudem, dass interdisziplinäre Crosskonkordanzen den stärksten (positiven) Effekt auf die Suchergebnisse haben.

Die Evaluation des Re-Ranking nach Bradfordizing (insgesamt 164 Fragestellungen) zeigt, dass die Dokumente der Kernzone (Kernzeitschriften) für die meisten Testreihen eine signifikant höhere Precision als Dokumente der Zone 2 und Zone 3 (Peripheriezeitschriften) ergeben. Sowohl für Zeitschriften als auch für Monographien kann dieser Relevanzvorteil nach Bradfordizing auf einer sehr breiten Basis von Themen und Fragestellungen an zwei unabhängigen Dokumentkorpora empirisch nachgewiesen werden.

Schlagwörter: Bradfordizing, Crosskonkordanzen, heterogen erschlossene Fachdatenbanken, Information Retrieval, Re-Ranking, Kernzeitschriften, Evaluation, Relevanz, digitale Bibliotheken, bibliographische Datenbanken, Zeitschriftenartikel, Monographien

Abstract

In spite of huge document sets for cross-database literature searches, academic users expect a high ratio of relevant and qualitative documents in result sets. It is particularly the order and structure of the listed results (ranking) that play an important role when designing search systems alongside the direct full text access for documents. Users also expect flexible information systems which allow influencing the ranking of documents and application of alternative ranking techniques.

This thesis proposes two value-added approaches for search systems which treat typical problems in searching scientific literature and seek to improve the retrieval situation on a measurable level. The two value-added services, semantic treatment of heterogeneity (the example of cross-concordances) and re-ranking on Bradfordizing, which are applied in different search phases, are described in detail and their effectiveness in typical subject-specific searches is evaluated in the empirical part of the thesis.

The preeminent goal of the thesis is to study if the proposed, alternative re-ranking approach Bradfordizing is operable in the domain of bibliographic databases, and if the approach is profitable, i.e. serves as a value added, for users in information systems.

We used topics and data from two evaluation projects (CLEF and KoMoHe) for the tests. The intellectually assessed documents come from seven academic abstracting and indexing databases representing social science, political science, economics, psychology and medicine.

The evaluation of the cross-concordances (82 topics altogether) shows that the retrieval results improve significantly for all cross-concordances, indicating that interdisciplinary cross-concordances have the strongest (positive) effect on the search results.

The evaluation of Bradfordizing re-ranking (164 topics altogether) shows that core zone (core journals) documents display significantly higher precision than was seen for documents in zone 2 and zone 3 (periphery journals) for most test series. This post-Bradfordizing relevance advantage can be demonstrated empirically across a very broad basis of topics and two independent document corpora as well for journals and monographs.

Keywords: Bradfordizing, cross-concordances, heterogeneous indexed databases, Information Retrieval, re-ranking, core journals, evaluation, relevance, digital libraries, bibliographic databases, journal articles, monographs

Widmung

Für meine Frau Kathrin und meine Söhne Quirin, Xaver, Vincent und Korbinian.
Für meine Eltern Irene und Wolf.

Danksagung

Ich bedanke mich bei allen Kolleginnen und Kollegen an den GESIS-Standorten Bonn und Berlin, die zum Gelingen dieser Arbeit beigetragen haben.

Meinen beiden Gutachtern gebührt besonderer Dank. Walther Umstätter und Jürgen Krause haben viele Stunden investiert, um aus vielen Ideen eine gute zu machen. Sie haben in unseren Gesprächen geholfen, diese Idee zu entwickeln, angemessen zu untersuchen und letztlich zu beschreiben. Ich habe viel von ihnen gelernt. Dafür danke ich ihnen beiden.

Der Arbeitsgruppe des KoMoHe-Projekts, bestehend aus Anne-Kathrin Walter, Vivien Petras, Max Stempfhuber, Stefan Bärish und Jürgen Krause, habe ich ebenfalls viel zu verdanken. Insbesondere meine beiden Kolleginnen Anne und Vivien haben mir durch ihre freundliche und kluge Art oft weitergeholfen und wissenschaftliche Probleme kleiner erscheinen lassen. Der Erfolg dieser Arbeit hat viel mit euch zu tun.

Meinem Kollegen Peter Mutschke will ich für seine besonnene Herangehensweise und die stets kritischen Anmerkungen danken.

Meinen drei Zimmernachbarn Philipp Schaer, Timo Wandhöfer und Werner Honnef danke ich, weil sie mich ein Jahr kaum gestört und meine geschlossene Tür akzeptiert haben. Die drei verbreiten zudem unglaublich gute Laune. Jetzt brechen endlich wieder andere Zeiten an.

Hanne Schott hat uns den Einstieg in das Projekt KoMoHe und die Untiefen der Heterogenitätsbehandlung (Crosskonkordanzen) sehr erleichtert. Guido Oemmelen und Matthias Stahl haben mir sehr kompetent mit den Signifikanztests geholfen. Stefan Bärish hat das in KoMoHe eingesetzte Assessment-Tool entwickelt und zur Verfügung gestellt. Dirk Hohmeister hat mir bei den Auswertungen der CLEF- und KoMoHe-Topics geholfen. Rielies Neitzke danke ich für die hervorragende Literaturversorgung. Bettina Zacharias danke ich sehr für ihre Mühe bei der Übertragung des Manuskripts in das Layout der GESIS-Schriftenreihe.

Frank Sonntag gebührt großer Dank für seine Ausdauer bei der akribischen Durchsicht dieses Manuskripts.

Ich danke allen und freue mich auf eine weitere gute Zusammenarbeit in der neuen GESIS.

1 Einleitung

Nutzer sehen sich je nach Art und Umfang der Literaturrecherche immer häufiger in der Situation, auf unterschiedliche Suchsysteme zugreifen zu müssen, um ihren Informationsbedarf zu befriedigen. Inzwischen führt die alleinige Nutzung z.B. eines Online-Katalogs (OPAC) der lokalen Bibliothek genauso wenig zu vollständigen Ergebnissen wie die Beschränkung bei der Suche auf eine einzelne Fachdatenbank. Ein Blick auf die aktuelle Informationslandschaft zeigt, dass viele Informationsangebote zwar dezentral online angeboten werden, aber zumeist unverbunden zur Verfügung stehen. So kann in einer Disziplin aus einer Vielzahl von unterschiedlichen singulären Literaturdatenbanken, Katalogen und anderen Informationssystemen gewählt werden, die bislang nicht an einer Stelle zusammengefasst bzw. recherchierbar sind. Zusätzlich bietet die Suche im frei zugänglichen Web über Internetsuchmaschinen weitere Möglichkeiten, um relevante Dokumente zu finden. Praktisch hat das zur Folge, dass mit jedem Wechsel zwischen den einzelnen Informationssystemen der Nutzer vor das Problem gestellt wird, die bereits für den OPAC formulierte Suche für eine oder mehrere Fachdatenbanken neu zu formulieren bzw. in weitere Systeme zu übersetzen. Dies erhöht die kognitive Belastung des Nutzers bei der Recherche und führt zwangsläufig zu längeren und unterbrochenen Recherchen. Häufig ist die Konsequenz, dass Nutzer aus Bequemlichkeitsgründen ausschließlich mit Internetsuchmaschinen recherchieren und die fachwissenschaftlich erschlossenen Literaturdatenbanken nicht mehr konsultiert werden.

Vor allem wissenschaftliche Nutzer erwarten heute möglichst die Bündelung und Konzentration fachlich relevanter und qualitativer Information an einer Stelle. Dadurch wird die Integration unterschiedlicher Fächer zur Vereinfachung des Zugangs bei interdisziplinären Fragestellungen notwendig. Die kontrollierten Vokabulare der Literaturdatenbanken, die in dieser Arbeit eine wichtige Rolle spielen, sind ein Hauptinstrument, um inhaltlich erschlossene Informationsangebote über Fächergrenzen hinweg zielgenau nach fachlichen Gesichtspunkten zu recherchieren. Bislang wissen vor allem erfahrene Rechercheure und Informationsspezialisten um die Möglichkeiten der kontrollierten Vokabulare, aber auch um die Problematik, dass im ungünstigsten Fall jede Datenbank die Dokumente mit einer anderen Indexierungssprache erschließt. Das Problem wird insbesondere dann virulent, wenn - wie gefordert - mehrere Datenbanken parallel über eine Suchanfrage recherchiert werden sollen. Aktuelle Standardisierungsaktivitäten im Bereich des Semantic Web zeigen, dass sich auch das W3C z.B. mit dem Standard SKOS bemüht, die kontrollierten Vokabulare mit XML-Technologien Semantic Web-fähig zu machen und damit auf das gleiche Ziel hinwirkt.

Trotz großer Dokumentmengen für datenbankübergreifende Literaturrecherchen erwarten Nutzer einen möglichst hohen Anteil an relevanten und qualitativen Dokumenten in den Trefferergebnissen. Insbesondere die Reihenfolge und Struktur der gelisteten Ergebnisse (Ranking) spielt, neben dem direkten Volltextzugriff auf die Dokumente, inzwischen eine entscheidende Rolle beim Design von Suchsystemen. Nutzer erwarten weiterhin flexible Informationssysteme, die es unter anderem zulassen, Einfluss auf das Ranking der Dokumente zu nehmen bzw. alternative Rankingverfahren zu verwenden.

In dieser Arbeit werden zwei Mehrwertverfahren für Suchsysteme vorgestellt, die die dargestellten Probleme ansatzweise lösen und damit die Situation bei der Recherche nach wissenschaftlicher Literatur messbar verbessern können. Die beiden Mehrwertdienste Crosskonkordanzen und Bradfordizing, die in unterschiedlichen Phasen der Suche zum Einsatz kommen und hier ausführlich beschrieben werden, werden im empirischen Teil der Arbeit bzgl. der Effektivität für typische fachbezogene Recherchen evaluiert.

1.1 Thematische Einordnung

Neben den klassischen Informationsanbietern Bibliothek, Fachinformation/Dokumentation und den Verlagen sind Internetsuchmaschinen inzwischen fester Bestandteil bei der Recherche nach wissenschaftlicher Information. Scirus (Elsevier, 2004) und Google Scholar sind zwei Beispiele für Suchdienste kommerzieller Suchmaschinen-Unternehmen, die eine Einschränkung auf den wissenschaftlichen Dokumentenraum anstreben und nennenswerte Dokumentzahlen in allen Disziplinen generieren. Der Vergleich der Treffermengen für beliebige Suchthemen zeigt, dass die Wahl des Suchsystems, des Dokumentenpools und der Dokumenttypen entscheidenden Einfluss auf die Relevanz und damit letztlich auch die Akzeptanz des Suchergebnisses hat.

Tabelle 1 verdeutlicht die Mengenunterschiede am Beispiel der Trefferergebnisse für die Suchbegriffe „search engines“ bzw. „Suchmaschinen“ in der allgemeinen Internetsuchmaschine Google, der wissenschaftlichen Suchmaschine Google Scholar (GS) und der größten fachübergreifenden bibliographischen Literaturdatenbank Web of Science (WoS). Der Anteil der Dokumente, die in diesem Fall eindeutig der Wissenschaft zuzuordnen sind (siehe GS und insbesondere WoS in Tabelle 1), liegt gegenüber der allgemeinen Websuche lediglich im Promille-Bereich. Dieses Beispiel veranschaulicht, dass es ausgesprochen problematisch sein kann, fachwissenschaftliche Fragestellungen ausschließlich mit Internetsuchmaschinen zu recherchieren. Der Anteil der fachwissenschaftlich relevanten Dokumente in diesem Trefferpool ist i. d. R. sehr gering. Damit sinkt die Wahrscheinlichkeit, wissenschaftlich Relevantes (z. B. einen Zeitschriftenaufsatz) auf den ersten Trefferseiten zu finden, deutlich ab.

Tabelle 1: Vergleich der Trefferzahlen von Google, Google Scholar (GS) und Web of Science (WoS) (abgefragt am 25.10.2007)

| Suchbegriffe | Google | GS | WoS | GS/Google in Promille | WoS/Google in Promille | WoS/GS in Promille |
|----------------|------------|---------|-------|--------------------------|---------------------------|-----------------------|
| search engines | 83.600.000 | 554.000 | 1.900 | 6,6 | 0,02 | 3,4 |
| Suchmaschinen | 16.200.000 | 7.410 | 0 | 0,5 | 0 | 0 |

Die drei oben genannten Suchsysteme (Google, GS und WoS) unterscheiden sich in mehrerlei Hinsicht fundamental und eignen sich daher gut, um in die Grundthematik dieser Arbeit einzuleiten.

Die obigen Suchsysteme erschließen zunächst unterschiedliche Suchräume, und dies auf sehr spezifische Weise. Während Google praktisch jede frei zugängliche und über Hyperlink adressierbare Information im Internet erfasst, gehen die beiden akademischen Suchsysteme deutlich selektiver bei der Inhaltserschließung vor. Google Scholar erfasst neben frei zugänglichen elektronischen Publikationstypen im Internet hauptsächlich wissenschaftliche Dokumente, die direkt von den akademischen Verlagen bezogen werden. Das WoS, das auf den unterschiedlichen bibliographischen Datenbanken und Zitationsindizes des ehemaligen „Institute for Scientific Information“ (ISI) basiert, selektiert gegenüber den rein automatischen brute-force-Ansätzen der Internetsuchmaschine über einen qualitativen Ansatz. In den Datenbanken des WoS werden ausschließlich internationale Fachzeitschriften erfasst, die ein kontrolliertes Peer-Review durchlaufen. Insgesamt werden ca. 8.000 Zeitschriften ausgewertet und über die Datenbank verfügbar gemacht.

Wie bereits erwähnt, spielt neben der Abgrenzung der Suchräume und Dokumenttypen die Zugänglichkeit und Relevanz der Dokumente eine entscheidende Bedeutung für den Benutzer. Die neueren technologischen Entwicklungen des Web Information Retrieval (IR), wie sie Google oder GS implementieren, werten insbesondere frei zugängliche Dokumente mit ihrer gesamten Text- und Linkinformation automatisch aus. Diese Verfahren sind vor allem deshalb erfolgreich, weil sie Ergebnislisten nach Relevanz gerankt darstellen, einfach und schnell zu recherchieren sind und direkt auf die Informationsquellen verweisen. Die qualitativen Verfahren der traditionellen Informationsanbieter (z. B. WoS) hingegen zeigen genau bei diesen Punkten (Ranking, Einfachheit und Volltextzugriff) Schwächen, überzeugen aber vor allem durch ihre Stringenz, in diesem Fall die selektive Aufnahme von qualitätsgeprüften Dokumenten in das System und die inhaltliche Erschließung der Dokumente. Aufgrund der Beschaffenheit der Dokumente in klassischen Informationssystemen wie z. B. bibliographischen Fachdatenbanken greifen die Ranking-Verfahren des Web IR genauso wenig wie die herkömmlichen textstatistischen Information Retrieval-Verfahren. Neue Ansätze, wie z. B. **der Einsatz von Autorenzentralitätsmaßen** beim Ranking, aber auch die in dieser Arbeit vorgestellten Verfahren, zeigen, dass die spezifische Struktur der Literaturnachweise und die Qualität der Daten für alternative Mehrwertdienste gewinnbringend eingesetzt werden kann.

Abgegrenzt wird Ranking oder Relevance Ranking von sog. Sortierungen z. B. nach dem Erscheinungsjahr der Publikation, obwohl hier die Grenze zu „nach inhaltlicher Relevanz“ gerankten Listen konzeptuell nicht sauber zu ziehen ist. Das Ranking von Dokumenten führt letztlich dazu, dass sich die Benutzer solcher Suchsysteme fokussiert mit den oberen Treffermengen (top-gerankt) eines Suchergebnisses beschäftigen. Der mittlere und untere Bereich eines Suchergebnisses wird häufig nicht mehr in Betracht bezogen. Bei großen Treffermengen wie den in Tabelle 1 **dargestellten (>1.000 Dokumente)** ist eine vollständige Rezeption eines Suchergebnisses ohnehin nicht mehr möglich. Auf Grund der Vielzahl an relevanten und verfügbaren Informationsquellen ist es daher notwendig, Kernbereiche in den Suchräumen zu identifizieren und diese anschließend dem Nutzer hervorgehoben zu präsentieren. Die Kernbereichsbildung kann zum einen über statistische Ranking-Verfahren erfolgen, zum anderen aber auf Grundlage der den Suchräumen inhärenten Strukturen. Das in dieser Arbeit evaluierte Verfahren Brad-

fordizing setzt auf einer typischen informatrischen Regelhaftigkeit thematischer Dokumentenräume auf und nutzt diese Regelhaftigkeit zum Re-Ranking von Dokumenten.

1.2 Motivation und Ziele

Die in dieser Arbeit vorgestellten Mehrwertverfahren a) Suchraumerweiterung durch Crosskonkordanzen (siehe Kapitel 3) und b) Re-Ranking durch Bradfordizing (siehe Kapitel 4) setzen an unterschiedlichen Stellen im Suchprozess ein und können zunächst singulär angewandt werden. Während die Suchraumerweiterung insbesondere zum Zeitpunkt der Suche aktiviert werden kann, setzt das Re-Ranking durch Bradfordizing nach Abschluss einer erfolgreichen Suche ein. Ziel dieser Arbeit ist es, die beiden Verfahren unabhängig voneinander zu beschreiben und über IR-Tests zu evaluieren.

Es wird angenommen, dass die Crosskonkordanzen als Verfahren der Suchraumerweiterung die Treffermenge einer Suchanfrage erhöhen. Der Nutzer erweitert dabei durch Hinzufügen semantisch relevanter Terme den Suchraum und wird durch dieses Verfahren i. d. R. deutlich mehr Dokumente generieren, insbesondere bei der Suche über mehrere verteilte Datenbanken. Der Einsatz von Crosskonkordanzen kann neben der Suchraumerweiterung aber auch zu einer Präzisierung der Suchanfrage und damit des Suchraums führen (mehr semantisch relevante Terme). Der Zugewinn an Dokumenten kann in jedem Fall erst dann als ein Mehrwert angesehen werden, wenn unter den hinzugewonnenen Dokumenten auch ein gewisser Anteil an relevanten Dokumenten in Bezug auf die Ausgangsfragestellung vorhanden ist. Diese Fragestellung soll in der Arbeit exploriert und evaluiert werden.

Der zweite Mehrwertdienst Re-Ranking durch Bradfordizing zielt auf die Neu- bzw. Reorganisation eines vorliegenden Trefferergebnisses und setzt damit konzeptuell nach der Suchraumerweiterung an. Es kann davon ausgegangen werden, dass die Anwendung von Bradfordizing nach dem Einsatz der Crosskonkordanzen zur Verstärkung der Suchverbesserung führen kann. Dem Bradford Law of Scattering (BLS), auf dem Bradfordizing basiert, liegt zugrunde, dass sich die Literatur zu einem beliebigen Fachgebiet bzw. -thema in Zonen unterschiedlicher Dokumentenkonzentration verteilt. Dem Kernbereich mit hoher Konzentration der Literatur folgen Bereiche mit mittlerer und geringer Konzentration. Bradfordizing sortiert bzw. rankt eine Dokumentmenge somit nach Kernzeitschriften. In dieser Arbeit soll überprüft werden, ob sich diese automatisch identifizierbaren Kernbereiche der Forschungsliteratur für den Nutzer zur Informationsrecherche eignen. Es wird angenommen, dass die Kernzonen/Kernzeitschriften nach Bradfordizing (hohe Artikelhäufigkeit zu einem Thema) positive Auswirkungen für den Nutzer zeigen. Es soll in unterschiedlichen Fachgebieten untersucht werden, ob Bradfordizing als Ergänzung der klassischen Ranking-Verfahren bei der Suche in bibliographischen Datenbanken eingesetzt werden kann. Bradfordizing kann im weitesten Sinne auch als Kompensationsverfahren für die erweiterten Suchräume des vorherigen Mehrwertdienstes beschrieben werden. Im Fokus dieser Arbeit liegt aber die Evaluation der Einzelverfahren.

Vorrangiges Ziel der Promotion ist es, zu untersuchen, ob das hier vorgestellte alternative Re-Rankingverfahren Bradfordizing im Anwendungsbereich bibliographischer Datenbanken zum einen operabel ist und zum anderen voraussichtlich gewinnbringend in Informationssystemen eingesetzt und dem Nutzer angeboten werden kann.

Im folgenden Abschnitt werden die Forschungsfragen dieser Arbeit aufgeführt und knapp erläutert. Eine ausführliche Beschreibung dieser Fragestellung und das entwickelte Untersuchungsdesign finden sich in Kapitel 5.

1.3 Forschungsfragen

Nachfolgend werden die zentralen Fragestellungen dieser Arbeit dargestellt. Die Reihenfolge der Fragestellungen entspricht im Prinzip der Abfolge der Experimente der einzelnen Untersuchungen.

1. *Welche Auswirkungen/Effekte hat das Konzept der semantischen Heterogenitätsbehandlung am Beispiel von Crosskonkordanzen für die Suche in Literaturdatenbanken?*
Es wird angenommen, dass eine Suche über verteilte, unterschiedlich erschlossene Datenbanken, die über die Verfahren der Heterogenitätsbehandlung, in dieser Arbeit der Mehrwertdienst Crosskonkordanzen (siehe Kapitel 3.2.1), erweitert wird, positive Effekte für die Suche in inhaltlich erschlossenen Dokumenten liefert. Der dazu konzipierte IR-Test hat das Ziel, die Anwendung von Crosskonkordanzen in einem realen Suchszenario zu evaluieren. In diesem IR-Test soll beantwortet werden, ob der Einsatz der Crosskonkordanzen die Suche gegenüber einer „nichttransformierten“ (nicht erweiterten) fachlichen Suche messbar verbessern kann.
2. *Inwieweit lassen sich Bradford-Verteilungen in unterschiedlichen Fachgebieten und insbesondere bei fachlichen Recherchen in Literaturdatenbanken identifizieren?*
Informetrische Verteilungen treten konstant und in unterschiedlichsten Konstellationen, vor allem aber in fachlich abgegrenzten Bibliographien auf. Die zweite Fragestellung untersucht dahingehend, ob sich die Gesetzmäßigkeit des Bradford Law in den analysierten Fächern und Datenbanken beobachten lässt. In dieser Studie wird Bradfordizing auf Dokumentenpools aus unterschiedlichen Fachdatenbanken aus den Bereichen Sozial- und Politikwissenschaften, Wirtschaftswissenschaften, Psychologie und Medizin zu verschiedenen Fragestellungen angewendet. Ziel der Untersuchung ist es, eine möglichst umfangreiche Anzahl von standardisierten Fragestellungen in unterschiedlichen Datenbanken und Fachgebieten durch Bradfordizing zu analysieren, um anschließend das Vorkommen der Bradford-typischen Verteilungen verifizieren oder falsifizieren zu können.
3. *Lassen sich informatrische Verfahren (z. B. Bradfordizing) als Mehrwertdienste für die Informationssuche in Literaturdatenbanken gewinnbringend nutzen?*
Die Anwendung informatrischer und vergleichbarer Verfahren ergibt eine alternative Sicht auf Ergebnismengen, die vom Benutzer i. d. R. als intuitiv und positiv empfunden werden. Das Bradfordizing einer Ergebnismenge liefert ein komplett neu- bzw. reorganisiertes Suchergebnis. Anhand eines qualitativen Interviews soll der Neuigkeitswert bzw. Mehrwert des Verfahrens Bradfordizing untersucht werden.

4. *Welche Auswirkungen/Effekte hat die Anwendung von Bradfordizing als Re-Ranking-Ansatz für das Information Retrieval? Lassen sich die Effekte innerhalb eines IR-Tests evaluieren?*

Es wird angenommen, dass die Anwendung von Bradfordizing zum Re-Ranking für ein thematisches Suchergebnis positive Effekte für die Suche in inhaltlich erschlossenen Dokumenten generiert. Bei der Untersuchung dieser Fragestellung stellen sich folgende Fragen: Kann Bradfordizing ganz generell als ein Mehrwertdienst in Informationssystemen (z. B. Fachdatenbanken) eingesetzt werden? Verbessert das Bradfordizing den Anteil der relevanten Dokumente im ersten Drittel (Core) einer Dokumentmenge gegenüber dem zweiten und dritten Drittel (Zone 2 bzw. Zone 3)? Ziel der Untersuchung ist es, den Re-Ranking-Ansatz Bradfordizing für Zeitschriftenartikel anzuwenden und in einem realen Suchszenario zu evaluieren.

5. *Lässt sich das Verfahren Bradfordizing als Re-Ranking-Ansatz auch auf andere Dokumenttypen und Szenarien übertragen? Kann Bradfordizing auch auf Monographien angewendet werden?*

Die letzte Fragestellung prüft, ob sich das Bradfordizing, das ursprünglich nur für Zeitschriftenliteratur beschrieben wurde, auch auf andere Dokumenttypen anwenden lässt. Für diesen Test werden die Relevanzbewertungen der Monographien zu den Fragestellungen der vorigen IR-Tests analysiert.

1.4 Aufbau der Arbeit

In Kapitel 2 werden die Informationstypen und Informationsangebote, die in dieser Arbeit eine Rolle spielen, am Beispiel von vascoda beschrieben. Die Darstellung konzentriert sich auf Literatur bzw. Literaturnachweise, insbesondere Zeitschriftenartikel und Monographien, sowie deren unterschiedliche inhaltliche Erschließungsmodelle (bibliothekarische und dokumentarische Erschließung). Ein Abschnitt widmet sich den elektronischen Publikationen und den Problemen, die bei der Integration dieses Publikationstyps in übergreifende Suchsysteme entstehen. Abschließend wird die akademische Suchmaschine Google Scholar beschrieben und empirisch untersucht.

In Kapitel 3 wird das Problem der semantischen Heterogenität zwischen inhaltlichen Erschließungssprachen bei der verteilten Suche über mehrere Fachdatenbanken beschrieben. Das Schalenmodell (Krause, 1996) und die daraus abgeleiteten Verfahren zur Heterogenitätsbehandlung (insbesondere das Konzept der intellektuell erstellten Crosskonkordanzen) werden als Lösungsansätze präsentiert. Die Ergebnisse des Forschungsprojekts KoMoHe, das ein umfangreiches Netz an Crosskonkordanzen erstellt und evaluiert hat, werden vorgestellt.

In Kapitel 4 wird die informatrische Gesetzmäßigkeit Bradford's Law of Scattering beschrieben. Das daraus abgeleitete Verfahren Bradfordizing und die Einsatzmöglichkeiten von Bradfordizing als Re-Ranking-Ansatz in bibliographischen Datenbanken werden ebenfalls in diesem Kapitel vorgestellt.

Kapitel 5 „Material und Methode“ beschreibt die in dieser Arbeit durchgeführten Untersuchungen. Es werden die Evaluationsmethoden, die Fragestellungen und Daten

der einzelnen Untersuchungen ausgeführt. Für die Tests wurden Fragestellungen und Daten aus zwei umfangreichen Evaluationsprojekten verwendet. Die evaluierten Dokumente stammen aus insgesamt sieben wissenschaftlichen Fachdatenbanken der Fächer Sozialwissenschaften, Politikwissenschaft, Wirtschaftswissenschaften, Psychologie und Medizin.

In Kapitel 6 werden einzelne Vorstudien präsentiert, die explorative Vorarbeiten zur Durchführung der Hauptuntersuchung darstellen. Die Vorstudien erkunden Fragestellungen, die nicht integraler Bestandteil der Hauptuntersuchung sind, die aber Hinweise auf die Anschlussmöglichkeiten der in dieser Studie vorgestellten Verfahren liefern. Die Vorstudien verdeutlichen Mehrwerte und Auswirkungen der in dieser Arbeit untersuchten Dienste.

Kapitel 7 präsentiert die Ergebnisse der in Kapitel 5 vorgestellten Untersuchungen. Das Ergebniskapitel führt die Ergebnisse der eingangs formulierten Forschungsfragestellungen zusammen. Der Schwerpunkt der Ergebnisdarstellung liegt bei der Evaluation des Mehrwertdienstes Bradfordizing.

In Kapitel 8 werden die Ergebnisse diskutiert. Die Diskussion der Ergebnisse zieht die Forschungsliteratur heran und erläutert den Neuigkeitsgewinn der Ergebnisse dieser Arbeit.

Kapitel 9 zeigt die anschließenden Forschungsfragestellungen auf. Dabei wird besonderer Wert auf die Kombination unterschiedlicher Mehrwertdienste gelegt.

In Kapitel 10 werden die Ergebnisthesen dieser Arbeit zusammengeführt.

2 Bibliographische Datenbanken, Portale und Suchmaschinen

In diesem Kapitel soll die aktuelle Situation bei der Zusammenführung von heterogenen wissenschaftlichen Dokumentkollektionen beschrieben und thematisiert werden. Die Abschnitte dieses Kapitels konzentrieren sich auf die Beschreibung der Bestandteile verteilter elektronischer Informationssysteme und deren Eigenschaften. Hintergrund der Darstellung ist die Situation, die für die Fachinformation im deutschsprachigen Raum am Beispiel von vascoda festgestellt werden kann. Die Überlegungen und Argumente stammen z. T. aus Arbeitsberichten innerhalb des Projekts KoMoHe (Krause und Mayr, 2006, Mayr, 2006), das in Abschnitt 3.2 dargestellt wird.

Ein ambitioniertes Projekt für die Suche nach wissenschaftlicher Literatur in Deutschland ist das Portal vascoda¹, das gemeinsam vom Bundesministerium für Bildung und Forschung und der Deutschen Forschungsgemeinschaft gefördert wird. vascoda bietet eine gemeinsame Suchoberfläche für eine Vielzahl von disziplinären und interdisziplinären Datenbanken und Sammlungen von Internetquellen (siehe dazu Abschnitt 2.4). Basiskonzept des vascoda-Portals ist die Strukturierung und Integration von intellektuell erschlossenen Informationsangeboten von mehr als 40 Datenlieferanten innerhalb eines Suchraums. Es werden insgesamt schätzungsweise 81 Millionen Dokumente nachgewiesen (Zeitpunkt Dezember 2007). Der Suchraum setzt sich aus verteilten disziplinären Portalen („Virtuellen Fachbibliotheken“ und Fachportalen) zusammen. Jede der integrierten Kollektionen wird in die vascoda-Fachgruppen eingeordnet (siehe dazu Abbildung 2-1). Das fachübergreifende vascoda-Portal enthält viele Informationsangebote, die sehr individuell entwickelt und strukturiert sind. Diese Kollektionen besitzen spezifische Metadatenschemata zur Beschreibung und Organisation der Inhalte. Seit 2007 ist vascoda Partner im globalen Wissenschaftsportal WorldWideScience.org.

Aus der Anfangsphase der vascoda-Initiative stammt folgendes Zitat (vgl. dazu auch (Depping, 2007) und aktuelles Gutachten zur Weiterentwicklung von vascoda von (Wolff, 2007)):

„Das Ziel von vascoda ist ein Dienst, der alle Bedürfnisse wissenschaftlich arbeitender Menschen befriedigt - von der Suche nach Literaturnachweisen bis zur Beschaffung des gewünschten Volltextes in elektronischer oder gedruckter Form sowie von dem Zugriff auf hochwertige Internetquellen. Dabei geht es vor allem darum, einen nach außen hin homogenen Zugriff auf verteilt vorliegende, heterogene Quellen zu ermöglichen.“ (Pianos, 2005)

Die übergreifende Suchoberfläche des Portals bietet jedoch nur eine Freitextsuche, die die ursprünglich präzisen Möglichkeiten des Fachzugangs dieser Informationskollektionen nicht berücksichtigt. Aufgrund des verteilten Aufbaus der Kollektionen und der unterschiedlichen Schemata der inhaltlichen Erschließung wird es als komplexes technisches und organisatorisches Problem gesehen, alle diese heterogenen Informationsangebote mit ihren individuellen und detaillierten Fachzugängen innerhalb einer gemeinsamen Suchoberfläche abzubilden.

1 <http://www.vascoda.de>

Auf der anderen Seite werden aktuell die großen Internet-Suchmaschinen um neue Kollektionen und erweiterte Möglichkeiten des Fachzugangs ergänzt (z. B. erweiterte Bilder- und Videosuche in Flickr² oder die zusätzlichen Funktionalitäten der alternativen Schnittstellen der Google-Suchmaschine³). Ein weiteres prominentes Beispiel sind die Semantic Web-Anwendungen⁴, die Deduktionsfunktionen durch den Einsatz von Ontologien und anderen semantischen Daten ableiten (vgl. (Krause, 2008)). Wie kann es möglich sein, dass die erweiterten Oberflächen für Digitale Bibliotheken in der Entwicklung zurückfallen, wenn moderne Web-Informationssysteme (wie z. B. aus der Semantic Web Initiative resultierend) sich gerade bemühen, mehr Struktur und semantische Klarheit zu schaffen?

Die Vision eines deutschen Wissenschaftsportals *vascoda* klingt im Vergleich mit den zuvor geschilderten Einschränkungen bei der fachübergreifenden Recherche sehr ambitioniert. Der geforderte ‚homogene Zugriff‘ auf heterogene, verteilte Quellen bereitet aus vielerlei Gründen (sowohl konzeptuell-technisch als auch rechtlich) große Schwierigkeiten. Das Beispiel des Allgemeinzugangs im Abschnitt 2.2, aber auch die Untersuchung der Google Scholar-Suchmaschine (siehe Abschnitt 2.7) verdeutlichen weitere Schwierigkeiten dieses geforderten homogenen Zugangs.

Welche konzeptuellen Probleme sich bzgl. der Heterogenität der erschlossenen Bestände ergeben, wird in dieser Arbeit beschrieben (siehe Kapitel 3). Insbesondere die Heterogenität der inhaltlichen Erschließung (semantische Heterogenität), die bei der Zusammenführung von unterschiedlich indextierten Quellen entsteht und dem ‚homogenen Zugriff‘ (z. B. einheitliche(s) Browsing, Deskriptorensuche) konzeptuell entgegensteht, soll hier thematisiert werden. Beispielsweise existieren neben den Unterschieden in der Erschließungs- bzw. Indexierungssprache (vgl. (Hodge, 2000, Lancaster, 1972)) auch Unterschiede in der Erschließungstiefe und damit der Qualität des schlagwortgebundenen Fachzugangs. Zu einer pragmatischen und groben Einteilung der in der Praxis vorkommenden kontrollierten Vokabulare bzw. Knowledge Organization Systems (KOS) kommt Hodge. Sie unterscheidet folgende Klassen von Vokabulartypen, die sie weiter unterteilt:

“The descriptions given here provide an overview of possible systems for organizing digital libraries. The descriptions are based on characteristics such as structure and complexity, relationships among terms, and historical function. The list is not comprehensive, nor are the definitions of these terms contained in specific standards documents. They are grouped into three general categories: term lists, which emphasize lists of terms often with definitions; classifications and categories, which emphasize the creation of subject sets; and relationship lists, which emphasize the connections between terms and concepts.” (Hodge, 2000)

Neben qualitativ hochwertig erschlossenen Informationsangeboten werden weitere z. T. kaum inhaltlich erschlossene Bestände in wissenschaftliche Fachportale einge-

2 <http://www.flickr.com/>

3 <http://code.google.com/>

4 Siehe Übersicht <http://www.w3.org/2001/sw/>

gliedert und auf der nächsthöheren Abstraktionsstufe *vascoda* zusammengeführt (vgl. dazu Konzept Schalenmodell in Abschnitt 3.1). In einem übergreifenden und umfassenden System mit steigender Anzahl der Teilangebote entsteht somit zwangsläufig eine zunehmend heterogene und komplexe Informationsarchitektur. Die Bandbreite von wissenschaftlichen Informationsangeboten im Fall *vascoda* umfasst die unterschiedlichsten Informationstypen und reicht von der Internetquelle über die klassischen Literaturnachweise bis hin zu Fakten, Volltexten und weiteren digitalen Materialien (siehe dazu Abschnitt 2.3). Insbesondere die beiden Erschließungstraditionen: Erschließung der Bibliotheken und Erschließung der IuD, die beide große Dokumentenmengen hinter sich vereinigen und sich bzgl. Tiefe und Spezifität deutlich unterscheiden, machen bei der Zusammenführung der Bestände (siehe dazu (Seiffert, 2001)) innerhalb *vascoda* eine wissenschaftlich begründete Modellbildung notwendig.

2.1 Ausgangssituation

vascoda definierte sich bis Ende 2006 auf oberster Ebene vor allem durch einen fachorientierten Zugang, durch eine exact match-Suche (Boolesche Algebra) und das Prinzip der verteilten Suche (auch Metasuche oder „distributed search“ genannt). Diese Grundidee des fachorientierten Zugangs für ein interdisziplinäres System wird im Gutachten von Chr. Wolff (Wolff, 2007) explizit als überzeugend und klar bewertet. Das Gutachten spricht allerdings auch problematische Punkte der aktuellen Realisierung des Fachzugangs an. So wird die Intransparenz und Komplexität der Informationsarchitektur sowie das Fehlen von „use cases“ angesprochen.

In den folgenden beiden Abschnitten werden Grundentscheidungen angesprochen, die exemplarisch für die Integration von heterogenen Beständen innerhalb eines Suchsystems stehen (siehe dazu insbesondere (Krause und Mayr, 2006)).

Konzeptuell sieht *vascoda* bisher vor, Fach-Kataloge bzw. Fach-OPACs zu bilden, die auf der Ebene der Fachportale integriert und kontrolliert werden. Eine Frage der Informationsarchitektur, die auch im Rahmen dieser Arbeit von Interesse ist, war, ob dieses fachorientierte Konzept als alleiniges Zugriffsmuster erhalten bleibt oder daneben ein zweiter Weg des (weniger tief erschlossenen) Allgemeinzugangs in fachübergreifenden Bibliothekskatalogen eröffnet wird, der den rein fachlich orientierten Zugang ergänzt. Wichtig wird diese Überlegung im Hinblick auf den Nutzer (siehe z. B. (Boekhorst, 2003)) vor allem dann, wenn man davon ausgeht, dass Nichtwissenschaftler wie z. B. Journalisten oder Schüler eine andere begriffliche Sicht auf einen Suchraum haben, die nicht nur gröbere Klassierungen voraussetzt, sondern konzeptuell unterschiedlich ist. Der Laie empfindet beispielsweise andere Klassenbildungen als „natürlich“ als der Wissenschaftler. Eine metaphorische Parallele hierzu findet sich in der Informatik, wenn auf dieselbe Datenstruktur verschiedene „views“ (Sichtweisen) zugelassen werden, um unterschiedlichen Nutzergruppen und Verwendungskontexten gerecht zu werden. Hinzu kommt, dass in der Regel allgemeine Recherchen (z. B. Ein-Wort-Anfragen) in fachübergreifenden

Bibliothekskatalogen (wie z. B. dem WorldCat⁵) im Vergleich zu kleineren Fachkollektionen mit Abstand die meisten Suchtreffer liefern (vgl. (Borgman, 1986, Borgman, 1996)).

Der vascoda-Suchdienst ist momentan auf der obersten Ebene in fünf Fachzugänge bzw. -cluster („Ingenieur- und Naturwissenschaften“, „Medizin und Biowissenschaften“, „Recht, Wirtschaft und Sozialwissenschaften“, „Geisteswissenschaften, Regionen und Kulturräume“ und „Fachübergreifende Kollektionen“) und einen fachübergreifenden Dokumentenpool eingeteilt (siehe Abbildung 2-1).

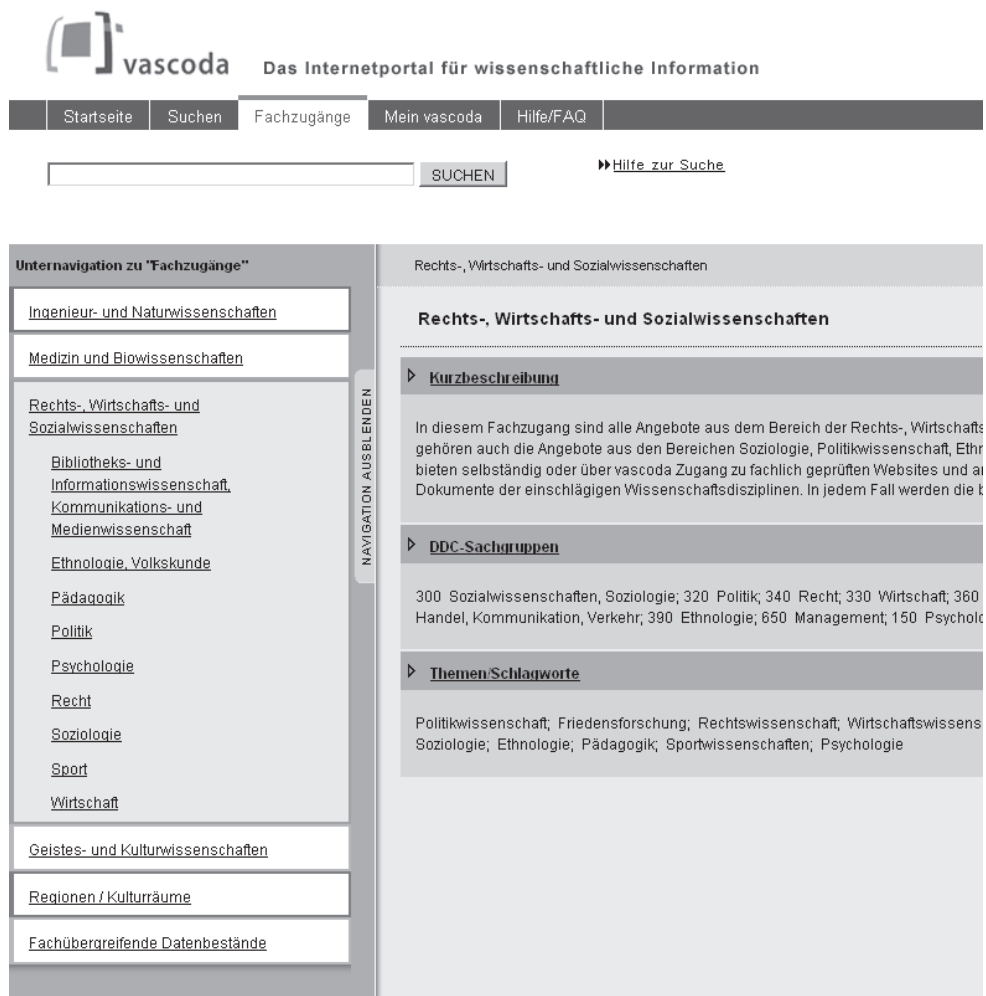


Abbildung 2-1: vascoda-Fachzugänge (Stand 2008). Fachzugang „Recht, Wirtschaft und Sozialwissenschaften“ sowie die untergliederten Fächer.

5 <http://www.worldcat.org/>

Hinter jedem dieser Fachzugänge befinden sich aktuell einzelne Fächer, die durch entsprechende Virtuelle Fachbibliotheken oder Fachportale repräsentiert werden. Die „fachübergreifenden Datenbestände“ spielen eine Sonderrolle, die im folgenden Abschnitt eingehend diskutiert werden soll.

2.2 Allgemeinsuche versus Fachzugang

Der Auslöser für die Diskussion Allgemeinsuche versus Fachzugang in (Krause und Mayr, 2006) war die Integration des umfangreichen Dreiländerkatalogs (DLK)⁶ in den Suchraum von vascoda.

Das ursprüngliche vascoda-Konzept geht davon aus, dass alle Fachportale aus den Angeboten der Verbundkataloge einen Fach-OPAC bilden, der dann in die Fachsuche integriert wird. Die Bibliotheksbestände waren damit als fachlich gefilterte Untermengen bei den Fachinformationen verortet, die dann über die Metasuche für Benutzer ohne Fachinteresse über alle Fächer hinweg zusammengeführt wurden (siehe dazu Ansatz bei (Chernov et al., 2006)). Folgende unterschiedliche Typen der Einbindung von OPACs finden sich in der Realität der Fachportale:

- 1) Fachlicher Auszug eines OPACs: Typischerweise wird über ein semi-automatisches Verfahren ein bestimmter fachlich eindeutig zuordbarer Bereich im Dokumentenbestand des Katalogs identifiziert. Dieser fachliche Ausschnitt kann bei Bedarf über eine Feldkennung separat abgefragt oder extrahiert werden. Beispielsweise kann die fachliche Selektion anhand der inhaltlichen Erschließung der Bestände über Klassifikationsbereiche erfolgen (z. B. OPAC der USB Köln, sozial- bzw. wirtschaftswissenschaftlicher Ausschnitt z. T. über Notationen der Basisklassifikation).
- 2) Fach-OPAC (gesamt): Spezialkataloge wissenschaftlicher Bibliotheken können aufgrund des Sammelschwerpunktes der Bibliothek direkt ohne vorherige Filterung in ein Fachportal integriert werden. Mit hoher Wahrscheinlichkeit ist der überwiegende Teil der Dokumente fachlich relevant (z. B. OPAC des Wissenschaftszentrums für Sozialforschung in Berlin für die Sozialwissenschaften).
- 3) Gesamter OPAC: Beispielsweise wird der gesamte Bibliothekskatalog einer Universitätsbibliothek in ein Fachangebot eingegliedert. Dies geschieht z. T. ohne eine fachliche Filterung der Bestände. Da es sich aber häufig um fachübergreifende Bestände handelt, kann von einer fachlich abgestimmten Recherche bei diesem Eingliedertyp nicht mehr gesprochen werden.
- 4) Verbundkatalog: Bei der Einbindung von Verbunddatenbanken wie z. B. der des GBV bzw. des HBZ ist mit einer Verstärkung der unter 3) aufgezeigten Tendenz zu rechnen, zumal auch Stadtbibliotheken und andere Bibliotheken mit eher generellem Sammelauftrag in den Verbund katalogisieren.

6 Der Dreiländerkatalog enthielt über 57 Millionen Titel (inkl. Dubletten, Stand Anfang Oktober 2006). Inzwischen existiert dieser Katalog nicht mehr. Ein Teil der Daten des Dreiländerkatalogs kann aber nach wie vor noch über das vascoda-Portal im „fachübergreifenden Datenbestand“ der Suchmaschine recherchiert werden.

Da die Suche im sog. fachübergreifenden Pool ausschließlich über das Termmaterial und nicht direkt über eine Aufteilung nach Fächern erfolgt, ergibt sich für allgemeine Anfragen eine wesentlich größere Zielmenge gegenüber den einzelnen Fachkollektionen. Der Hauptgrund für die deutlich größeren Treffermengen im fachübergreifenden Pool sind fachgebietsübergreifende Termidentitäten, die nicht vorab ausgefiltert werden.

Ob die Integration von fachübergreifenden Verbundkatalogen anstelle von inhaltlich geprüften Fach-OPACs als Übergangslösung für interdisziplinäre Portale akzeptabel ist, müssen empirische Tests mit Relevanzbewertung zeigen. Eine Abweichung vom fachorientierten Prinzip legt beim Benutzer in jedem Fall eine Konzeptänderung nahe: Der fachübergreifende Pool ließe sich als „allgemeiner Bibliothekszugang“ bzw. Allgemeinzugang interpretieren, zusätzlich zu den weiter bestehen bleibenden Fachzugängen.

Die Literaturübersicht⁷ im folgenden Abschnitt setzt sich mit der Problematik Allgemeinsuche versus Fachzugang auseinander. Ziel ist es, die Eigenschaften der unterschiedlichen Erschließungstraditionen auf Basis der Literatur zu charakterisieren und damit mehr Klarheit zu den Fragen der Zusammenführung dieser Bestände zu schaffen.

2.2.1 Analytische Literatúrauswertung

Immer wieder finden sich in der wissenschaftlichen Literatur und in den Förderprogrammen von DFG, EU und BMBF des letzten Jahrzehnts Forderungen nach einer Zusammenführung der Fachdatenbanken mit den Bibliothekskatalogen. So formuliert z. B. das BMBF in seinem strategischen Positionspapier „Information vernetzen – Wissen aktivieren“ explizit die Zusammenführung von Bibliothekszugängen und Fachinformationsangeboten als Ziel und fordert dazu auf, für mehr Konvergenz zwischen den Informationsbeständen über einen übergreifenden Ansatz zu sorgen.

„Die traditionelle Informationsinfrastruktur aus Bibliotheken und Fachinformationszentren muss in leistungsfähige vernetzte Organisationssysteme transformiert werden.“ (BMBF, 2002) S. 4)

Allgemein geteilte Begründungen hierfür finden sich in Konzepten wie dem „one stop shop“ in vascoda (vgl. vascoda-Papier „Strategie vascoda“). Grundlegend ist, dass Benutzer an Antworten interessiert sind und nicht an organisatorischen Strukturen und Hintergründen der Informationserschließung. Wenn somit die Unterschiede zwischen den einzelnen Informationssammlungen in den verschiedenen institutionellen Kontexten Bibliotheken, IuD-Stellen oder Universitäten im heuristisch-pragmatischen Sinn für die Unterschiede der Informationssammlungen verantwortlich sind, kann man sich ohne weitere Überlegungen der Aufgabe widmen, die bestehende Heterogenität durch geeignete konzeptuelle und pragmatische Maßnahmen aufzuheben (siehe dazu Abschnitt 3.2.1). Das Ergebnis wäre ein Informationspool, dessen Informationsangebote die Differenzierung nach Bibliothekskatalog und IuD-Stelle hinter sich lassen.

Anders anzugehen ist die Zusammenführung, wenn hinter den Informationssammlungen von Bibliotheken und denen der IuD-Stellen konzeptuelle Unterschiede stehen, die

⁷ Der Abschnitt „Analytische Literaturbetrachtung“ ist in gekürzter Form dem Arbeitsbericht Krause und Mayr (2006) entnommen.

auch in einer neuen Situation des integrierten Zugangs erhalten bleiben sollen. Ist somit aus der Tradition der Bibliotheken so etwas wie eine eigene Sichtweise auf die Welt der wissenschaftlichen Literatur ableitbar, die hier als „Allgemeinzugang“ oder „Allgemeinsuche“ bezeichnet werden soll – im Gegensatz zur fachbezogenen Konzeptualisierung der wissenschaftlichen Welt, wie sie am besten in Fachontologien und Fachthesauri⁸ abgebildet ist?

Folgt man dieser Überlegung, wäre zunächst zu fragen, worin sich der Allgemeinzugang vom Fachzugang konzeptuell unterscheidet, d. h., wenn man heuristische und pragmatische Argumente beiseite lässt.

Der entscheidende inhaltliche Unterschied der beiden Zugangstypen liegt in der fachlichen Selektierbarkeit der Dokumente. Während man bei einer fachlich erschlossenen Literaturdatenbank davon ausgehen kann, dass hier vornehmlich fachrelevante Zeitschriftenliteratur (siehe (Umstätter, 2003)) und Monographien nachgewiesen werden, kann eine vergleichbare fachlich eindeutige Zugriffsmöglichkeit für den Allgemeinzugang nicht in jedem Fall gewährleistet werden. Aber das reicht nicht für eine konzeptuelle Trennung von Fach- und Allgemeinzugang. Die fachliche Differenzierung könnte beispielsweise als einfacher Klassifikationsfilter im Rahmen eines einheitlichen Zugangs realisiert werden. Erst wenn damit eine andere Sichtweise auf die wissenschaftliche Fachliteratur verbunden ist, die für Benutzer außerhalb der Fachwelt oder bei interdisziplinären Recherchen „natürlicher“ als Zugang gewählt wird, macht es Sinn, von einem konzeptuellen Unterschied auszugehen. Metaphorisch vergleichbar hiermit wäre in der Informatik das Datenbankkonzept verschiedener „views“ (Sichtweisen) auf dieselbe zugrunde liegende Datenstruktur.

Als ein Ansatzpunkt für solch eine Sichtweise lässt sich F. Geißelmann ((Geißelmann, 1999) S. 41) interpretieren:

„Die bibliothekarische Erschließung ist eine *fachübergreifende Erschließung* [Hervorhebung des Autors]. Die Notwendigkeit einer fachübergreifenden Suche wird trotz aller Betonung der Fachinformation sogar von den wissenschaftlichen Fachgesellschaften anerkannt. Dass bibliothekarische Sacherschließung kein Gegensatz zur dokumentarischen sein muss, sondern beide aufeinander abgestimmt werden sollten, ist unter diesen Gesichtspunkten eine Forderung der Zukunft.“

Es macht misstrauisch, dass sich in der Literatur keine konkreten Hinweise zu einer konzeptuellen Notwendigkeit eines Allgemeinzugangs finden lassen, der sich zweifelsfrei als Begründung einer „view“-Änderung interpretieren ließe. Der „Geißelmann-Forderung“ lässt sich – zumindest in einer der möglichen Lesarten des Zitats – durch die Realisierung von Suchräumen über alle Fächer hinweg nachkommen.

8 Die klassische Definition eines Thesaurus findet sich in Umstätter, Walther und Wagner-Döbler, Roland (2005): Einführung in die Katalogkunde: Vom Zettelkatalog zur Suchmaschine, 3., vollständig neu überarbeitete Auflage, Hierseemann, Stuttgart. „Ein Thesaurus im Bereich Bibliothek und Dokumentation ist nach DIN 1463 Teil 1 eine „geordnete Zusammenstellung von Begriffen und ihren (vorwiegend natürlichsprachigen) Bezeichnungen, die in einem Dokumentationsgebiet zum Indexieren, Speichern und Wiederauffinden dient“.

Die in der Literatur angegebenen Gründe für Unterschiede zwischen Fachinformation und bibliothekarischer Information lassen sich in der Regel heuristisch aus dem Entstehungszusammenhang heraus interpretieren.

Die Ziele bibliothekarischer Erschließung sind traditionell etwas anders verortet und wurden im Prinzip bereits bei Cutter definiert (hypostatization of users' needs) (zitiert nach (Svenonius, 2000) S. 15).

„According to Cutter, those objectives were

1. to enable a person to find a book of which either
the author
the title
the subject is known
2. to show what the library has
by a given author
on a given subject
in a given kind of literature
3. to assist in the choice of a book
as to its editions (bibliographically)
as to its character (literary or topical).

Cutter formulated his objectives based on what the user needs and has in hand when coming to a catalog.”

Diese klassischen drei Ziele von Bibliothekskatalogen und deren Erschließung werden 1. als finding objective, 2. collocating objective (Zusammenstellung von Dokumenten) und 3. als choice objective benannt. Sie wurden 1997 von der IFLA reformuliert ((Svenonius, 2000) S. 16ff.) und stellen bis heute die Basis für alle üblichen Katalogsysteme dar.

Es ist im Folgenden durchaus klar, dass in der Realität eine große Vielfalt an Bibliothekstypen und Bibliothekserzeugnissen (insb. Kataloge) für sehr unterschiedliche Nutzergruppen existiert, dass man schwerlich von der Bibliothek, dem Bibliothekskatalog bzw. der bibliothekarischen Erschließung sprechen kann. Die Bandbreite reicht von Öffentlichen Bibliotheken mit sehr allgemeinem Sammel- und Erschließungsauftrag bis hin zu Spezialbibliotheken und Forschungsbibliotheken mit fachlich sehr ausgefeilten Erschließungssystemen (vgl. die Neuauflage der Katalogkunde (Umstätter und Wagner-Döbler, 2005)). Die folgenden Argumente stellen daher sehr generelle Aussagen dar:

- Bibliotheken wenden sich an ein anderes Publikum mit der Konsequenz, dass Unterschiede bei den aufgenommenen Informationstypen und Inhalten entstehen, die in den nichtwissenschaftlichen Bereich verweisen. „Sacherschließung in Bibliotheken diene bislang in erster Linie dazu, Sach- und Fachbücher zu erschließen. Sie wendet sich im Gegensatz zur Dokumentation meist an ein *breites Publikum* [Hervorhebung des Autors] bzw. an diejenigen, deren Fragen erfahrungsgemäß in der Bibliothek beantwortbar sind, also Fragen zu bestimmten Personen bzw. deren Publikationen, Fragen zur Geographie oder zur Geschichte, oder auch zu bestimmten Sachthemen. (...) Wird also ein spezieller Gipfel der Ostalpen in einem Dokument beschrieben, so kann der eine Thesaurus in seiner Indexierungsspezifität alle

diese Gipfel einzeln enthalten, ein zweiter mag bei den Ostalpen, ein dritter bei den Alpen und ein vierter bei der Benennung Gebirge enden. Dabei spielt natürlich auch die Indexierungsbreite eine Rolle, die im Bibliotheksbereich meist so gering ist, daß man froh sein kann, wenn überhaupt noch ein zweites oder drittes Sachschlagwort vergeben wurde. Allein diese Tatsache verbietet schon das angestrebte enge Schlagwort für Bibliotheken unterschiedlicher Größe.“ (Umstätter, 1991)⁹ Wie der DLK bzw. „fachübergreifende Pool“ von vascoda als Prototyp eines Allgemeinzugangs sehr gut zeigt (gleiches gilt auch für den WorldCat), befinden sich sehr unterschiedliche Dokumenttypen in den zusammengeführten Bibliotheksbeständen (z. B. Monographien, mehrbändige Werke, Hochschulschriften, Zeitschriften, Dissertationen, Aufsätze, Konferenzbände, Statistiken, Berichte, Multimedia-Objekte, Serien, Festschriften, elektronische Ressourcen, amtliche Veröffentlichungen, Mikroformen usw.). Die Vielfalt der Dokumenttypen wird dabei durch eine Reihe von Dokumenttypen ergänzt, die nicht zu den wissenschaftlichen Nachweisen zu zählen sind und ausschließlich von Bibliotheken nachgewiesen werden (z. B. Schöne Literatur, Kinder- und Jugendliteratur, Schulbücher, Musiktonträger, Musikvideos usw.). In der Regel kann der Anteil der fachlich nicht zuweisbaren Dokumente innerhalb eines Katalogs nicht ohne Weiteres bestimmt werden. Demnach kann nicht vermieden werden, dass nichtwissenschaftliche Nachweisinformationen (z. B. Schöne Literatur, Schulbücher) ausgegeben werden. Der Übergang ist hier aber sicher fließend. Für Literaturwissenschaftler sind Primärtexte Schöner Literatur vor allem mit textkritischem Apparat sicher relevant.

Solange es um die Nachfrage nach wissenschaftlicher Information geht, sollte ein Portal auch den Bedürfnissen z. B. von Journalisten oder Schülern gerecht werden. Da die bibliothekarische Inhaltserschließung auf dieses breitere Publikum ausgerichtet ist, müsste der Zugang zumindest über die bibliothekarischen Erschließungssysteme wie die Schlagwortnormdatei (SWD) und die Universalklassifikationen (z. B. Basisklassifikation) ausreichen.

- Bibliotheken haben einen einfacheren Zugang zur Originalliteratur mit der Konsequenz einer geringeren Anforderung an die Inhaltserschließung.
„Er [der Bibliothekskatalog, Anmerkung des Autors] grenzt sich insofern von dokumentarischen Datenbanken bzw. Bibliographien ab, als diese nicht das Ziel der Bestandserschließung einer bestimmten Bibliothek haben, sondern eine weitergehende Erschließung bestimmter Fachgebiete, wie z. B. Anatomie, Chemie, Geschichte, Medizin, Philosophie und Soziologie.“ ((Umstätter und Wagner-Döbler, 2005) S. 11)
- Bibliotheken legen mehr Wert auf die Formalerschließung.
„Bei der Erschließung ergibt sich traditionsgemäß eine größere Betonung auf die Formalerschließung, da die zu erschließenden Medien im bibliothekarischen Kontext meist umfangreicher sind als im dokumentarischen.“ ((Hobohm, 2004) S. 512)

9 Siehe dazu die Definitionen für Indexierungsspezifität, Indexierungsbreite bei Ewert, Gisela und Umstätter, Walther (1997): Lehrbuch der Bibliotheksverwaltung, Hiersemann, Stuttgart, ISBN: 3-7772-9730-5.

(Taylor, 1999) betont bei der bibliothekarischen Erschließung den Vorrang der 'authority work', was sich am ehesten mit formaler Erschließung im Sinne einer Normierung der „Ansetzungen“ übersetzen lässt, und mit der finding objective (indikative Erschließung¹⁰) in Deckung gebracht werden kann. Normierungsbemühungen wie z. B. das aktuelle Projekt Virtual International Authority File (VIAF) der DNB, LoC und OCLC zeigen, dass in übergreifend und kooperativ erschlossenen Datenbeständen auf konsistente (quasi „normalisierte“) Ansetzungen über Normdaten nicht mehr verzichtet werden kann (siehe (Bennett et al., 2006)).

„Cataloging of individual items involves creating a description of the physical item; choosing certain names and titles to serve as access points for getting to the description in the catalog; doing authority work on those names and titles; doing subject analysis of the content of the work in the item; choosing subject headings and classification numbers to represent the subject analysis ...“ ((Taylor, 1999) S. 6) vascode strebt an, über eine Vereinheitlichung der Metadaten zur Formal- und Sacherschließung bestehende Unterschiede zumindest in einem Kernbereich aufzuheben. Auch wenn man von einer Differenzierung nach dem Umfang der Informationsquelle ausgehen würde (z. B. Bücher und deren stärkere Formalerschließung), hätte dies keinen Einfluss auf den hier diskutierten konzeptuellen Unterschied zwischen Allgemein- und Fachzugang.

- Die Wissenschaftler stellen höhere Anforderungen an die Erschließung und die aufzunehmenden Informationstypen, die von den IuD-Stellen erfüllt werden. Vor allem empfinden Wissenschaftler das Fehlen der unselbständigen Literatur in den Bibliothekskatalogen als unzureichend¹¹.

„Während man im Bereich der Bibliothek im 18. und 19. Jh. die inhaltliche Erschließung vornehmlich durch bibliothekarische Klassifikationen zur Ordnung der Bücher verwendete, ergänzte im 20. Jh. die Dokumentation die gezielte Auffindbarmachung auch unselbständiger Publikationen durch ihre erheblich tiefer gehende Erschließung. Damit wurde dem weiter wachsenden Aufkommen an Literatur Rechnung getragen. Im 21. Jh. setzt sich diese Entwicklung einer immer feineren Erschließung nun durch die Organisation einzelner Wissens Elemente in Wissensbanken fort.“ (Ewert und Umstätter, 1997)

Gründe für die Abtrennung der Information und Dokumentation (IuD) vom Bibliothekswesen finden sich beispielsweise in (Seeger, 2004) S. 183ff; siehe dazu auch (Umstätter, 2005):

„Einbeziehung sog. grauer Literatur, Wirtschafts- und Patentinformation, differenzierte Auswertung von Zeitschriftenaufsätzen etc.“

„Notwendigkeit der differenzierten Inhaltsererschließung der Dokumente und deren Auszügen, Zusammenfassungen“.

¹⁰ Vgl. dazu den Begriff der „informativen Erschließung“ in der IuD.

¹¹ Nichtsdestotrotz wird die Erarbeitung von Bibliographien oder heute Fachdatenbanken durchaus als bibliothekarische Aufgabe gesehen.

Bibliothekare erschließen heute in der Regel mit genormten und umfangreichen Vokabularen oder Klassifikationen. Die Indexierung von Zeitschriftenaufsätzen wird wiederum mit spezifischen Thesauri und anderen kontrollierten Vokabularen vorgenommen. Die größere Breite der Erschließung erklärt sich beispielsweise, wenn man die Anzahl der durchschnittlich vergebenen Deskriptoren mit dem Umfang des verschlagworteten Dokuments vergleicht. Eine Monographie mit durchschnittlich 150 Seiten erhält bei bibliothekarischer Verschlagwortung selten mehr als 3 Deskriptoren. Die Praxis der IuD zeigt, dass ein Zeitschriftenartikel mit durchschnittlich 10 Seiten mit ca. 10 Deskriptoren indexiert wird. Die Aufsatzerschließung im bibliothekarischen Bereich („Primat des engen Schlagworts“) beläuft sich ebenfalls auf etwa 3 Deskriptoren pro Artikel und ändert somit am Verhältnis 3:10 nichts. Demnach erschließen Bibliotheken vorwiegend indikativ und synoptisch¹², IuD-Stellen dagegen informierend, zusammenfassend (vgl. Abstracting).

Konzeptuell entscheidend ist nicht die Frage, ob einzelne Informationstypen tiefer zu erschließen sind oder erschlossen werden, sondern die Frage, ob die größere Erschließung eine andere Sichtweise auf die Welt der wissenschaftlichen Inhalte widerspiegelt. Ist das nicht der Fall, gäbe es keinen speziellen Allgemeinzugang im Sinne der „view“-Metaphorik. Ob der Benutzer feiner oder gröber anfragen möchte, würde sich nur in den Heterogenitätskomponenten widerspiegeln, die unterschiedliche „real existierende“ Erschließungsdifferenzen ausgleichen (siehe dazu Kapitel 3.2.1).

Gegen die Ansetzung eines Allgemeinzuganges spricht, dass bibliothekarische Klassifikationen und Thesauri in der Regel die Untergliederung der Fachwissenschaften als Basis nehmen¹³. Diese fachwissenschaftliche Untergliederung lässt sich an den unterschiedlich tief ausgearbeiteten Universalklassifikationen, aber auch dem so genannten Universal-Thesaurus SWD (Sachgruppen-Struktur) bzw. den Library of Congress Subject Headings (LCSH) ablesen. Bei diesen universell angelegten Beschreibungssprachen wird versucht, die Breite der Terminologie der Wissenschaften bzw. der zu erschließenden Literatur möglichst gut abzudecken. Da das Deskriptorenmaterial beispielsweise der SWD über die Sachgruppen fachlich gegliedert ist, wäre es in Teilbereichen durchaus operabel, Fachauszüge aus dem „fachübergreifenden Pool“ von vascoda zu erstellen und separat anzubieten. Damit würde anstatt des Konzepts Allgemeinzugang das fachlich ausgefeiltere und mit vascoda kompatible Konzept der Fach-OPACs umsetzbar. Es ergeben sich auch aus dieser Betrachtung keine Hinweise auf die Notwendigkeit, einen Allgemeinzugang als konzeptuelle Differenzierung dem vascoda-Portal zugrunde zu legen.

12 Siehe dazu die Definition von Bibliothek in Ewert, Gisela und Umstätter, Walther (1997): Lehrbuch der Bibliotheksverwaltung, Hiersemann, Stuttgart, ISBN: 3-7772-9730-5.

13 Dies ist natürlich kein „Beweis“, dass nicht doch ein Allgemeinzugang im Sinne der view-Metapher von Nutzen wäre. Es ist nur ein Hinweis darauf, dass es in der Literatur und Praxis bisher keine Verfechter solch eines Modells zu geben scheint.

2.2.2 Fazit

Es finden sich in der Literatur unserer Kenntnis nach keine Hinweise, die auf eine abgekoppelte spezifisch bibliothekarische Sicht der bekannten universellen Beschreibungssprachen hindeuten.

Eine Änderung der bisherigen vascoda-Grundkonzeption (fachorientierter Zugang) lässt sich aus der Literaturübersicht nicht stützen. Es gibt einige wichtige pragmatische Gründe für Unterschiede in den Benutzergruppen, den Inhalten und der Breite und Tiefe der Inhaltserschließung. Es gibt jedoch keine Hinweise, dass sich hieraus die Notwendigkeit einer unterschiedlichen Sichtweise auf die Welt ergibt, die sich nicht als Differenzierung der Granularität der Informationsangebote interpretieren lässt. Damit reichen die Modelle der Heterogenitätskomponenten und das Schalenmodell als Grundkonzeption aus (siehe dazu Kapitel 3).

Nicht vergessen werden darf, dass mit der vorgeschlagenen Begründung nur wissenschaftlich-rationelle Begründungen für ein Modell geliefert werden, das sich im empirischen Test bewähren muss (siehe dazu mehrere Beispiele in Kapitel 5). Diese Aussage gilt generell für alle Modelle in der Softwareergonomie, Mensch-Maschine-Interaktion und dem Information Retrieval.

Der folgende Abschnitt löst sich von der Diskussion Allgemeinzugang vs. Fachzugang und beschreibt konkrete Informationstypen und -angebote, die sich auf den unterschiedlichen Abstraktionsebenen des vascoda-Portals befinden.

2.3 Informationstypen

Die Unterscheidung der Informationsangebote im vascoda-Umfeld in Literatur, Internetquellen (z. B. Informationen aus und über die Community und WWW-Dokumente) und Primärinformation ist eine pragmatische Einteilung, die helfen soll, die Vielfalt der Informationsangebote und -typen im vascoda-Umfeld zu bündeln (siehe dazu Beispiel in Abbildung 2-2). Es wird im Folgenden versucht, allgemeine Aussagen zum Stand der Erschließung der drei abstrakten Informationstypen zu treffen. Knappe Charakterisierungen sollen hier ausreichen.

2.3.1 Literatur

Bei dem Informationstyp Literatur handelt es sich um qualitätsgeprüfte publizierte Fachinformation in Form von Literaturnachweisen und digitalen Volltexten. Der größer werdende Bereich der erschlossenen „grauen Literatur“ (Hochschulschriften, graue Online-Literatur, Preprints, Reports usw.; siehe dazu auch Abschnitt 2.6), der von fast allen vascoda-Anbietern für die Recherche zur Verfügung gestellt wird, soll ebenfalls zu dem Typ Literatur gezählt werden. Literatur wird in der Regel formal und inhaltlich nach festgelegten Regeln erschlossen. Die formale oder bibliografische Erschließung in Deutschland z. B. nach den „Regeln für die alphabetische Katalogisierung“ (RAK) erfasst dabei die wichtigsten Angaben einer Literaturquelle (z. B. Autor, Titel, Verlag, ISBN, ISSN usw.) mit dem Ziel, die Quelle schnell und eindeutig wiederzufinden. Bibliothekare und

Dokumentare legen hier zwar unterschiedliche Kriterien an, grundsätzlich werden aber, bezogen auf den Literaturtyp, weitestgehend die gleichen Daten erfasst. Die inhaltliche Erschließung von Literatur, aber auch anderen Informationstypen bereitet sowohl beim Erschließen (z. B. Indexierungskonsistenz) als auch bei der Recherche einige Schwierigkeiten. So variieren die kontrollierten Vokabulare¹⁴ zur inhaltlichen Erschließung bezogen auf den Umfang, die Spezifität und die zu erschließende Literaturart (siehe dazu insbesondere Kapitel 3).

Abbildung 2-2: Erweiterte Suche des Fachportals sowiport (Beispiel für selektierbare Informationstypen)

Meistens liegt der Informationstyp Literatur heute für die Recherche lediglich als Nachweisinformation (Metadaten) vor, in zunehmendem Maße aber auch als digitale Volltextinformation (siehe dazu Abschnitt 2.6). Der Wunsch des Nutzers, direkt in Volltexten zu recherchieren bzw. sie vom Arbeitsplatz z. B. über Dokumentliefersdienste aus zu beziehen, ist laut mehrerer Benutzerstudien unverkennbar (siehe z. B. (Wolff, 2007)).

2.3.2 Internetquellen

Eine Sonderrolle innerhalb vascoda spielt der Informationstyp Internetquellen. Hierunter fallen insbesondere die Fachinformationsführer oder auch Subject Gateways ((Koch, 2000) siehe Abschnitt 2.4.3). Dabei handelt es sich um formal nach Dublin Core¹⁵ er-

¹⁴ Siehe dazu das erweiterte Konzept der Knowledge Organization Systems (KOS) bei Hodge.

¹⁵ Siehe <http://dublincore.org/>

geschlossene Internetseiten, die zusätzlich inhaltlich über kontrollierte Vokabulare erschlossen werden. Sowohl das Dublin Core Element `dc.type`¹⁶ als auch die angepassten oder selbst entwickelten Systematiken zur Ressourcenbeschreibung von Internetquellen bei den einzelnen Informationsanbietern sollen Struktur in die Vielzahl der Informationstypen bringen. Ein einheitliches und umfassendes Kategorienschema über alle angebotenen Informationstypen, das über die Empfehlung der vascoda-Arbeitsgruppe Metadaten hinausgeht (siehe vascoda Application Profile), ist aber bislang nicht sichtbar. Zudem werden fachbezogene Informationsangebote wie z. B. Experten- und Institutionsverzeichnisse, aber auch andere Formen der Information über und aus der wissenschaftlichen Community eines Fachportals (Veranstaltungsnachweise, Foren usw.) aufgebaut, die als zusätzliche Quellen zur Ergänzung für die Recherche eingesetzt werden können (siehe dazu Beispiele in Abbildung 2-2).

Webseiten und Homepages sind per se kaum oder gar nicht erschlossen. Erst über die Sammlungen mit Internetquellen im Kontext vascoda wird versucht, diesen Informationstyp für die Fachrecherche zu nutzen; allerdings lediglich über ihre formalen und inhaltlichen Metadaten. Ansätze der automatischen Indexierung, die den Volltext der Webseiten bzw. Webseitenstrukturen integrieren, sind erweiterte Einsatzmöglichkeiten dieses Informationstyps (vgl. (Lossau, 2004, Summann und Lossau, 2004)), die zunächst evaluiert werden müssen.

2.3.3 Primärinformation

Der deutlich seltener vertretene Informationstyp Primärinformation, in der Statistik z. T. auch Primärdaten genannt, umfasst all jene Information, die nicht das Ergebnis, sondern der Gegenstand oder das Hilfsmittel wissenschaftlicher Forschung sind (z. B. Umfragedaten¹⁷, Fakteninformation, Werkstoffdaten, Chemiedaten usw.). Primärinformationen sind zwar z. T. ebenfalls durch spezifische Klassifikationen erschlossen (siehe dazu Text-Fakten-Integration bei (Krause et al., 1997, Krause und Stempfhuber, 2005)), spielen aber zur Zeit für die Recherche in den Fachangeboten eine untergeordnete Rolle und sollen daher zunächst ausgeklammert werden. Internetdokumente können ebenfalls als Primärinformation angesehen werden und relevante Informationen und Daten beinhalten. Die automatische Eingliederung dieses Informationstyps in ein Wissenschaftsportal, das ansonsten ausschließlich kontrollierte und qualitätsgeprüfte Dokumente (hauptsächlich Literaturnachweise) anbietet, stellt momentan konzeptuell Probleme dar und soll an dieser Stelle nicht weiter ausgeführt werden.

2.3.4 Fazit

Die drei abstrakten Informationstypen Literatur, Internetquellen und Primärinformation können an dieser Stelle als ein pragmatisch-heuristischer Ansatz (kleinster gemeinsamer

¹⁶ Siehe <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>

¹⁷ Siehe dazu den Datenbestandskatalog der GESIS. <http://www.gesis.org/dienstleistungen/daten/recherche-datenzugang/datenbestandskatalog/>

Nenner) zur Kategorisierung und Vereinfachung der Informationsvielfalt innerhalb der vascoda-Angebote dienen. Eine Trennung der beiden Hauptklassen Literatur und Internetquellen erscheint sowohl auf der informationswissenschaftlich-konzeptuellen Ebene als auch für den Nutzer sinnvoll und plausibel.

Nachfolgend sollen die Informationsangebote, die die unterschiedlichen Informationstypen nachweisen, knapp charakterisiert werden.

2.4 Informationsangebote

Nachfolgend werden einzelne Informationsangebote, die in vascoda Eingang finden, charakterisiert und bezüglich ihrer Erschließung beschrieben (vgl. Bestandsaufnahme der Informationsangebote in vascoda in (Mayr, 2006)). Abbildung 2-3 zeigt die Vielfalt der Informationsangebote (im Kontext vascoda-Module genannt). In den folgenden Abschnitten wird eine Auswahl dieser Informationsangebote vorgestellt.

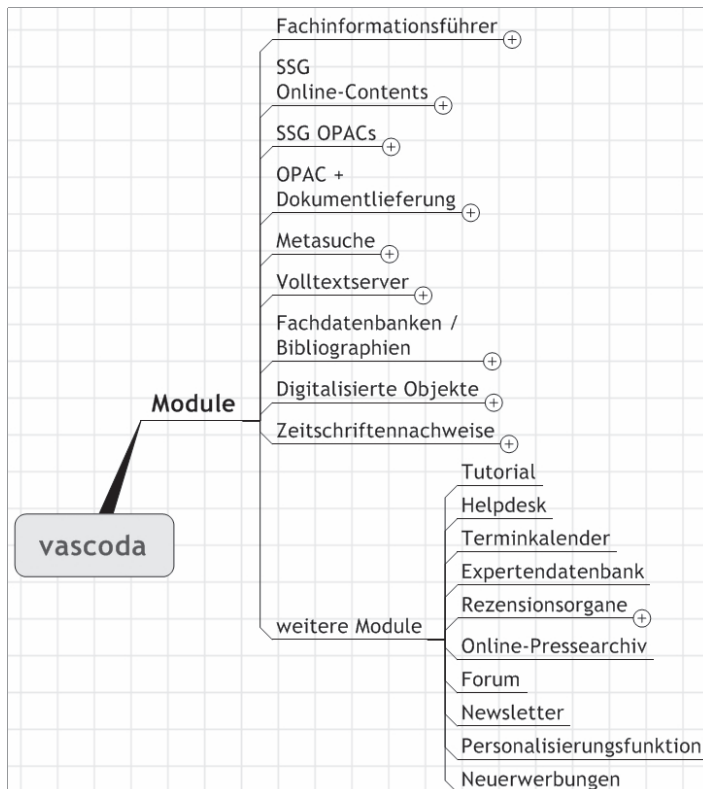


Abbildung 2-3: Informationsangebote im Kontext vascoda

2.4.1 Fachdatenbanken & Bibliographien

Der Bestand der Fachdatenbanken und Bibliographien, insbesondere der Informationsverbünde (siehe Abschnitt 2.5.2), ist sehr umfangreich und meist fachlich klar abgegrenzt. Die von der DFG finanzierten Nationallizenzen¹⁸ für wissenschaftliche Literatur, die Zugriff auf bibliographische Datenbanken, elektronische Zeitschriften und Volltexte bieten, haben das Spektrum der bundesweit frei zugänglichen Fachdatenbanken zudem deutlich erweitert. Die Größe der untersuchten Datenbanken variiert dabei deutlich (zwischen wenigen tausend Datensätzen bis zu mehreren Millionen wie z.B. der medizinischen Fachdatenbank Medline). Neben Monographien und Sammelwerken werden insbesondere einzelne Sammelwerksbeiträge sowie Artikel aus Fachzeitschriften und Proceedings ausgewertet und erfasst. Die Erschließung der Dokumente ist tiefgehend und wird in der Regel von Fachkräften (Indexierer) vorgenommen. Neben Autor und Titel werden meist detaillierte Informationen zur Quelle (z. B. Abstract) sowie weiterführende Angaben zum Dokument (wie etwa die Kontaktadresse des Autors) geliefert. Die inhaltliche Erschließung erfolgt über spezifische kontrollierte Vokabulare (Klassifikationen, Thesauri). Meist existiert pro Nachweis ein Abstract und vereinzelt auch Titelübersetzungen.

„Fachdatenbanken und Bibliographien sind eine sehr wichtige Informationsressource für die Virtuellen Fachbibliotheken, da sie (in der Regel per Autopsie) verifizierte Metadaten und nach einschlägigen fachlichen Standards inhaltlich erschlossene Angaben über das gesamte Schrifttum eines Faches bzw. Fachgebietes enthalten. Eine Fachdatenbank definiert sich inhaltlich darüber, dass sie ein Fachgebiet inklusive verwandter Bereiche der Nachbarfächer abdeckt (Scope).“ (vascoda, 2006)

2.4.2 OPACs

„Die Online-Kataloge weisen ... gedruckte Publikationen (Zeitschriften und Monographien) nach und ermöglichen über die Verknüpfung mit einer Dokumentlieferung den Zugriff auf den gewünschten Volltext, selbst wenn das Original nicht elektronisch vorliegt.“ (vascoda, 2006)

Neben einer einfachen Freitextsuche liefern die in Mayr untersuchten OPACs (Mayr, 2006) zusätzlich eine erweiterte Suche mit differenzierten Suchmöglichkeiten (siehe dazu (Bates, 1986, Dempsey, 2006, Eversberg, 2002, Larson et al., 1996)). Die Literaturbestände sind inhaltlich überwiegend mit der SWD und den RSWK erschlossen. Die formale Erfassung der Bestände erfolgt weitestgehend nach den RAK, daher überrascht es, dass die einzelnen analysierten OPACs die Abfragefelder sehr uneinheitlich für die Literaturrecherche zur Verfügung stellen. Praktisch alle OPACs der großen SSG-Bibliotheken und der Spezialbibliotheken werden durch Standardsoftware realisiert. Neben den fachlich selektierbaren Beständen der Universitäts-, Landes- und Stadtbibliotheken sind gerade für den Fachwissenschaftler die Bestände der Bibliotheken einzelner Forschungseinrichtungen (Spezialbibliotheken) von Interesse. So umfasst z.B. der OPAC des Wissenschaftszentrums Berlin für Sozialforschung (WZB) ca. 130.000 Medieneinheiten sozial-,

¹⁸ <http://www.nationallizenzen.de/>

politik- und wirtschaftswissenschaftlicher Fachliteratur, darunter mehr als 10.000 Publikationen aus der WZB-Forschung sowie ca. 430 laufend gehaltene Fachzeitschriften.

2.4.3 Sammlung von Internetquellen

„Fachinformationsführer oder subject gateways erschließen fachwissenschaftlich relevante - und in der Regel frei zugängliche - Internetressourcen (thematische Websites) auf der Grundlage standardisierter Metadaten (Dublin Core; Fach- und Universalklassifikation; Fach- oder Universalthesaurus). Die Tiefe der Erschließung einzelner thematischer Websites kann dabei von der Erstellung eines Metadatensatzes für eine Website reichen bis zur Erstellung von Metadatensätzen zu jedem einzelnen Dokument einer thematischen Website.“ (vascoda, 2006)

Die Mehrzahl der Quellensammlungen im Kontext vascoda weisen lediglich Internetstartseiten (Homepages von Institutionen usw.) und nur sehr vereinzelt Literatur nach. Der SocioGuide¹⁹ (siehe Abbildung 2-4) erschließt beispielsweise Institutionen, Netzwerke, Sammlungen, Veranstaltungen und Zeitschriften aus allen Teildisziplinen der Sozialwissenschaften und ausgewählten Themengebieten.

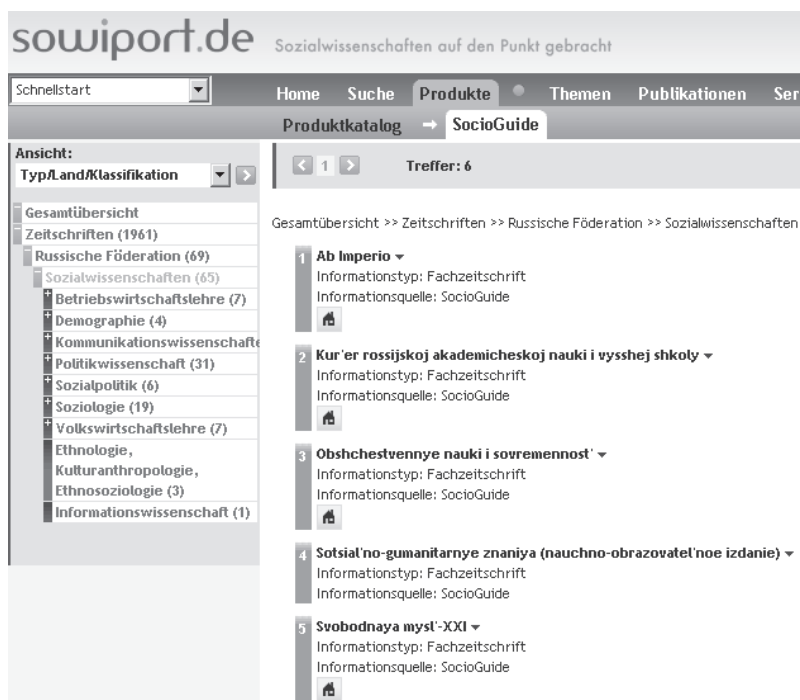


Abbildung 2-4: SocioGuide (Ausschnitt der russischen Fachzeitschriften aus dem Fachgebiet Sozialwissenschaften)

19 <http://www.gesis.org/dienstleistungen/fachinformationen/datenbanken-informationssysteme/socioguide/>

2.4.4 Dokumentenserver

„vascoda sichert den direkten Zugriff auf elektronische Volltexte und gewährleistet ihre langfristige Verfügbarkeit. Dies geschieht durch die Einrichtung von Volltextservern an den koordinierenden Bibliotheken der Fachportale, die Organisation des Zugriffs auf die Volltextserver nichtkommerzieller und kommerzieller Partner von vascoda bzw. auf den von vascoda in Zusammenarbeit mit der AG Verbünde zu realisierenden Verteilten Dokumentenserver.“ (vascoda, 2006)

Volltext-, Publikations- oder Dokumentenserver finden sich noch recht vereinzelt unter den Fachangeboten der Informationsanbieter. Laut Befragung der vascoda-Partner (März 2006) planen 10 Anbieter, künftig Dokumentenserver aufzubauen und mit anzubieten. Ein Beispiel für einen fachspezifischen Dokumentenserver ist z.B. das Social Science Open Access Repository (SSOAR)²⁰. Es zeigt sich, dass der Bereich der Open Access Repositories und Zeitschriften weltweit deutlich ansteigt, was die Verfügbarkeit von Volltextdokumenten verbessert. Das Directory of Open Access Repositories (DOAR²¹) listet z.B. mehrere hundert Dokumentenserver weltweit. Die Bielefeld Academic Search Engine (BASE²²) hat in den letzten Jahren einen umfangreichen Index aufgebaut, der Zugang zu über 16 Millionen Dokumenten bietet (Stand Oktober 2008), die sich auf den frei zugänglichen Dokumentenservern befinden.

2.4.5 Digitalisate

„Die Virtuellen Fachbibliotheken weisen retrodigitalisierte Bücher und Zeitschriften ihres Fachbereichs umfassend nach. Idealerweise wird dabei der Zugang bis auf die Einzeldokumentebene (bei Zeitschriften Aufsatzebene) erschlossen.“ (vascoda, 2006)

Digitalisierte Objekte (Digitalisate) finden sich noch seltener als Dokumentenserver unter den Fachangeboten der vascoda-Partner. Das Projekt DigiZeitschriften²³ kann beispielsweise als ein umfassendes Digitalisierungsprojekt von Fachzeitschriften im deutschsprachigen Raum angeführt werden. Die Inhalte von DigiZeitschriften werden von vascoda bislang nicht erfasst.

Nachdem die wichtigsten wissenschaftlich relevanten Informationsangebote vorgestellt worden sind, sollen im Anschluss die Informationsanbieter charakterisiert werden, die diese Informationsangebote zur Verfügung stellen.

20 <http://www.ssoar.info/>

21 <http://www.opendoar.org>

22 <http://www.base-search.net/>

23 <http://www.digizeitschriften.de/>

2.5 Informationsanbieter

Eine klare Zuordnung der oben erwähnten Informationstypen und -angebote zu einem bestimmten Anbietertyp ist auf Grund der technologischen Entwicklung bei den Anbietern und dem veränderten Sammelauftrag meist nicht möglich. Beispielsweise bieten sowohl Bibliotheken als auch die Fachinformationsanbieter der IuD „Literatur“ (z. B. Fachdatenbanken), aber auch „Internetquellen“ und andere Angebote an. In Kapitel 3 wird die Situation der Informationsversorgung aus Sicht eines Informationsanbieters dargestellt. J. Krause spricht in diesem Zusammenhang vom polyzentrischen Dokumentenraum (Krause, 2003), in dem viele Angebote dezentral über das WWW angeboten werden.

„So reicht die Spannbreite der Anbieter von nicht institutionell gebundenen Privatpersonen über öffentlich geförderte Institutionen unterschiedlichster Größe und Couleur bis hin zu kommerziellen Anbietern. Oft haben die unterschiedlichen Fächer eine bestimmte Präferenz für eine oder mehrere Anbietergruppen. Einzig die Bibliotheken und das WWW sind für praktisch alle Fächer von fundamentaler Bedeutung.“ (Müller, 2004)

Da die detaillierte Darstellung der Fachinformationsanbieter nicht im Mittelpunkt dieses Berichts steht, soll eine sehr knappe Typisierung der Anbieter bzgl. ihrer Informationsangebote ausreichen.

2.5.1 Bibliotheken

Bibliotheken (meist große Universitätsbibliotheken mit Sondersammelauftrag) sind ein klassischer Lieferant von kostenfreier Nachweisinformation und damit ein Kernangebot für die wissenschaftliche Informationsversorgung (Degkwitz, 2004, Hacker, 2005). Als Nachweis- und Rechercheinstrument werden meist OPACs bzw. Verbundkataloge (z. B. GBV, KOBV, Bibliotheksverbund Bayern, KVK) zur Verfügung gestellt, die Nachweise über Monographien, Sammelwerke, Graue Literatur, Dissertationen z. B. standardisiert liefern (siehe dazu auch Abschnitt 2.2).

Das Konzept der Virtuellen Fachbibliotheken entstammt der Förderlinie der DFG. Diese drittmittelgeförderten Projekte haben das Ziel, umfassend Bestände der Bibliotheken über ein portalähnliches System dem Fachwissenschaftler zugänglich zu machen und nachhaltig zu betreiben (siehe z. B. MedPilot²⁴, die Virtuelle Fachbibliothek Medizin). Insbesondere die integrierte Suche (Metasuche) über unterschiedliche Informationsbestände (beispielsweise verschiedene fachlich relevante OPACs, Fachdatenbanken) wird als Ziel definiert. Das Beispiel des Sammelauftrags der Deutschen Nationalbibliothek für Netzpublikationen²⁵ zeigt, dass sich das Spektrum der Informationsangebote laufend erweitert.

²⁴ <http://www.medpilot.de>

²⁵ <http://www.d-nb.de/netzpub/index.htm>

2.5.2 Fachinformationsanbieter

Die Informationsverbünde sind vom BMBF geförderte Projekte. Ziel ist ein Zusammenschluss von Fachinformationsanbietern, die fachgebietsrelevante Literatur (meist Zeitschriftenaufsätze und Sammelwerksbeiträge) über Fachthesauri erschließen und anbieten. Die Informationsverbünde bieten eine gemeinsame Oberfläche an, über die in fachlich relevanten und verwandten Datenbanken recherchiert werden kann. Zusätzlich besteht die Möglichkeit der angeschlossenen Dokumentlieferung. Beispiele für Informationsverbünde sind MedPilot, der Informationsverbund Bildung, Sozialwissenschaften, Psychologie infoconnex (siehe Abschnitt 3.2.2.1) und der Informationsverbund Naturwissenschaft und Technik getinfo²⁶. Die größten Anbieter von Literaturdatenbanken in Deutschland sind z. B. das Fachinformationszentrum Karlsruhe, das FIZ Technik sowie GENIOS.

2.5.3 Kommerzielle Anbieter

Verlage, kommerzielle Firmen (z. B. Datenbank-Hosts wie GENIOS), wissenschaftliche Fachgesellschaften (auch Eigenverlage), Dokumentlieferdienste sind häufig Lieferanten von kostenpflichtigen Volltext-Informationen (siehe (Williams, 2006)). Die Datenbanken können direkt über die Webangebote der Informationsanbieter oder indirekt über die Suchwerkzeuge der Informationsverbünde recherchiert werden. Inzwischen kooperieren einige große Internet-Suchmaschinen mit den kommerziellen und nichtkommerziellen Anbietern von Fachinformation und bieten spezielle Suchangebote zur Recherche in deren Beständen an. Prominentes aktuelles Beispiel ist Google mit seinem Suchdienst Google Scholar (siehe dazu Abschnitt 2.7).

2.5.4 Fachinformation im Internet

Fachinformation im Internet bezeichnet den stetig wachsenden Bereich der „grauen“ Onlineliteratur (z. B. Eprints, Preprints, Arbeitsberichte usw.), der qualitätsgeprüften Fachliteratur (z. B. Ejournal oder auch Open Access-Zeitschriften) und das weite Feld der WWW-Dokumente inkl. der neuen Publikationsformen wie z. B. Weblogs und Wikis, die Fachinformation enthalten, aber nicht kontrolliert erschlossen und in der Regel nur über Internet-Suchmaschinen zugänglich sind ((Lawrence und Giles, 1998, Lawrence und Giles, 1999), siehe dazu auch Abschnitt 2.6). Unterschiedliche Untersuchungen zeigen, dass diese Volltext-Informationsquellen für die wissenschaftliche Kommunikation immer relevanter und die freie Verfügbarkeit bedeutsamer werden (siehe Abschnitt 2.6.1). Die Kostenfreiheit und Offenheit des Zugangs spielt für den Nutzer hierbei eine große Rolle.

Besonderes Augenmerk soll daher im folgenden Abschnitt auf den elektronischen Publikationen liegen. Dieser spezielle Informationstyp spielt eine Sonderrolle, weil die Menge der rein elektronischen Veröffentlichungen stetig zunimmt, aber von den klassischen Informationsanbietern bei der Erschließung vernachlässigt wird (siehe dazu z. B.

26 <http://www.getinfo-doc.info/>

(McVeigh, 2004)). Dieser Abschnitt basiert auf einem Artikel des Autors im Sammelwerk „Wissenschaftsforschung Jahrbuch 2007“ (Mayr, 2007).

2.6 Elektronische Publikationen

Elektronische Publikationen umfassen ein breites Spektrum von Veröffentlichungstypen, die von Disziplin zu Disziplin sehr individuelle Ausprägungen besitzen (Brody, 2006, Nentwich, 2003). Betrachtet man beispielsweise unterschiedliche elektronische Publikationsformen, stellt man fest, dass die Grenzen zwischen einer klassischen Publikation - z. B. ein referierter Zeitschriftenartikel - und wissenschaftlichen Dokumenten im elektronischen Raum schnell verschwimmen. Es fehlen operable und allgemein etablierte Verfahren, um ein elektronisches Dokument eindeutig von einer elektronischen Publikation zu unterscheiden. Die Autopsie eines Dokuments alleine hilft zur Beurteilung jedenfalls meist nicht weiter.

Es stellt sich somit die Frage: Was ist eine elektronische Publikation? Was unterscheidet eine elektronische Publikation von einem Dokument, das elektronisch frei zugänglich und damit quasi „veröffentlicht“ ist?

Eine sehr einfache Definition, die sich allerdings nur auf das Medium der Verbreitung bzw. die Art der Zugänglichkeit bezieht, findet sich bei (Kling und McKim, 1999):

„We define an electronic publication as a document distributed primarily through electronic media.“ (S. 891)

Nach dieser Definition zählen alle Veröffentlichungen, die primär über Internet-Dienste wie Email oder WWW elektronisch zugänglich gemacht werden, zu den elektronischen Publikationen. Diese Definition setzt allerdings auf der Nutzerseite voraus, dass der Betrachter einer solchen Publikation die Herkunft und den Status des Dokuments beurteilen kann. Davon kann aber bei der Recherche im Internet nur in den seltensten Fällen ausgegangen werden. Der überwiegende Teil der wissenschaftlichen Dokumente, die über Suchmaschinen oder andere Suchsysteme zugänglich sind, sind dem Recherchierenden unbekannt und damit zunächst nicht eindeutig als Publikation zuzuordnen. Hinzu kommt, dass allgemeine Internet-Suchmaschinen (siehe z. B. (Brin und Page, 1998, Lewandowski, 2008)) nur zu einem sehr geringen Anteil wissenschaftliche Dokumente und zu einem noch geringeren Teil wissenschaftliche Publikationen nachweisen.

Somit stellt sich die Frage: Was ist eine Publikation? Bzw. im Kontext dieses Beitrags: Was ist eine wissenschaftliche Publikation?

Kling & McKim zählen elektronische Vorveröffentlichungen nicht automatisch zu den Publikationen, sondern benutzen dafür den neutralen Begriff ‚working article‘. Das entscheidende Kriterium, dass ein wissenschaftliches Dokument bzw. ein ‚working article‘ demnach nachweisen muss, ist das Peer-Review. Erst ein erfolgreich absolviertes Peer-Review macht aus einem wissenschaftlichen Dokument (z. B. Manuskript) eine Publikation (vgl. (Ginsparg, 2003)).

„We refer to articles made publicly and electronically available in nonpeer-reviewed form, either as posted on an individual or organizational Web page, or on a server such

as the Los Alamos Physics E-Print Archive (<http://www.arxiv.org/>), as electronic working articles.” ((Kling und McKim, 1999) S. 891-892)

Mit dieser Einschränkung auf ein erfolgreich durchlaufenen Peer-Review ist zwar ein formales Kriterium für die Unterscheidung von Publikationen und anderen wissenschaftlichen Dokumenten getroffen, hilfreich ist diese Unterscheidung aber hauptsächlich definitorisch. Nichtsdestotrotz wird diese Definition von vielen Wissenschaftlern und insbesondere den Verlagen und Informationsanbietern unterstützt und vehement verteidigt. Die Aufweichung dieser Definition würde auch schlichtweg bedeuten, dass alle elektronischen Dokumente, die im Internet „publiziert“ werden und auf Qualitätssicherungsverfahren wie das Peer-Review verzichten, ebenfalls zu den elektronischen Publikationen gehören würden. Jede x-beliebige Webseite, jeder Blog-Eintrag und jedes verschickte Email-Dokument mit wissenschaftlichem Inhalt wäre damit automatisch eine elektronische Publikation. Gegen diese Öffnung der Definition einer elektronischen Publikation sind aber selbst die Befürworter (Harnad, 1991, Harnad und Brody, 2004) des elektronischen Publizierens; folglich akzeptieren die meisten die Einschränkung, dass ein Dokument, gleich welchen Typs, erst eine Form von Qualitätssicherung durchlaufen haben muss, um als Publikation angesehen zu werden. Das heißt aber wiederum nicht, dass in wissenschaftlichen Publikationen keine Nicht-Publikationen zitiert werden können, was an sich zwangsläufig zu einer Erweiterung des Publikationsbegriffs führt. Letztlich müssen sich Autoren aber entscheiden, ob sie ihre Materialien elektronisch ohne Qualitätssicherung bereitstellen wollen oder ob sie den Weg des klassischen Publizierens gehen wollen (Posting vs. Publishing). Beide Wege werden in der Praxis beschritten und sind für sich genommen auch beide sinnvoll (siehe in diesem Zusammenhang den h-index für Autoren bei (Bornmann und Daniel, 2005)).

Wie eingangs bereits angeführt, hat sich in den letzten Jahren (Ginsparg, 2007) der Bereich der frei zugänglichen elektronischen Dokumente (Publikationen sowie andere Dokumententypen) in den meisten wissenschaftlichen Disziplinen deutlich erweitert. Open Access und Self-Archiving (Brody, 2006, Swan und Brown, 2005) sind die beiden bekanntesten Platzhalter für diese Entwicklung im elektronischen Publizieren. In zunehmendem Maße werden über dieses alternative Publikationsmodell entstandene Dokumente in Digitalen Bibliotheken in der Regel im Volltext verfügbar gemacht. Digitale Bibliotheken wie z. B. arXiv²⁷, CiteSeer²⁸ oder RePEc²⁹ spielen inzwischen innerhalb der jeweiligen Disziplinen eine wichtige Rolle bei der Informationsversorgung der Wissenschaftler. Eine Besonderheit dieser Digitalen Bibliotheken besteht darin, dass die Forschungsdokumente dort bereits vor der eigentlichen Veröffentlichung in Zeitschriften oder anderen Publikationsformen von den Autoren selbst zugreifbar gemacht werden und zusätzlich weitere Services bieten (siehe dazu (Chu und Krichel, 2003, Klas, 2007, Lawrence et al., 1999)). Verbreitet ist parallel dazu, dass Autoren ihre Manuskripte oder

27 <http://arxiv.org/>

28 <http://citeseer.ist.psu.edu/>

29 <http://repec.org/>

auch elektronischen Publikationen auf institutionellen oder privaten Internetseiten bzw. Homepages verfügbar machen, wo sie von Internet-Suchmaschinen gefunden werden.

2.6.1 Open Access & Self-archiving

Spätestens seit der Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen³⁰, in der die großen Forschungsorganisationen erklären, den freien Zugang zu Forschungsinformationen aktiv zu unterstützen, zeichnet sich ab, dass mittelfristig die Menge der Open Access (OA)-Publikationen stark zunehmen wird. Der Anteil der reinen OA-Publikationen, also elektronischen Publikationen, die in einer OA-Zeitschrift publiziert werden, am gesamten Publikationsaufkommen ist aktuell aber noch relativ gering (McVeigh, 2004). Dafür lassen sich unterschiedliche Gründe anführen (siehe dazu auch Überblick in (Stempfhuber, 2008)):

- mangelnde Bekanntheit und Nutzung der Möglichkeiten, OA zu publizieren (Fournier, 2005),
- mangelndes Vertrauen in OA,
- Verhaftung und Zufriedenheit mit dem traditionellen Publikationsmodell.

Was lässt sich gesichert über Open Access aussagen?

(Lawrence, 2001) hat als einer der ersten darauf hingewiesen, dass die durchschnittliche Anzahl der Zitationen für online verfügbare Artikel im Bereich der Informatik deutlich über den Artikeln liegt, die nicht online, also nur offline als klassische Print-Publikation zugänglich sind.

„The mean number of citations to offline articles is 2.74, and the mean number of citations to online articles is 7.03, an increase of 157 %.” (Lawrence, 2001)

Diese Tatsache wird sehr häufig als unbestreitbarer und eindeutiger OA-Effekt ((Harnad und Brody, 2004, Harnad et al., 2004)) interpretiert. Dass die aktuelle Situation von OA nicht derart eindimensional betrachtet werden kann, zeigen aktuelle Untersuchungen (z. B. (Kurtz et al., 2005)). OA befördert demnach vor allem die Einfachheit des Zugriffs auf die Publikation und kann generell die wissenschaftliche Kommunikation beschleunigen.

“... has the potential to accelerate recognition and dissemination of research findings, but its actual effects are controversial (...) Articles published as an immediate OA article on the journal site have higher impact than self-archived or otherwise openly accessible OA articles (...) OA articles are more immediately recognized and cited by peers than non-OA articles” (Eysenbach, 2006)

„Die Eprints überrunden die Zeitschriftenaufsätze in dem Sinne, dass 3/4 von ihnen in Eprints von anderen Autoren zitiert werden, bevor das entsprechende Heft von Physical Review D erscheint.“ (Havemann, 2004)

Des Weiteren liegen dem OA-Publizieren noch relativ wenig untersuchte Phänomene wie z. B. das Phänomen der Verzerrung durch einen ‚quality bias‘ für OA publizierte

30 http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf

Dokumente zugrunde. Qualitativ höher zu bewertende Publikationen werden dabei von den Autoren häufiger OA publiziert als deren durchschnittlichere Arbeiten.

“... evidence of a strong quality bias and early view effect. No sign of a general ‘open access advantage’ (...) ArXiv accelerates citation, due to the fact that ArXiv makes papers earlier available rather than that it makes papers freely available“ (Moed, 2006)

„We demonstrate conclusively that there is no “Open Access Advantage” for papers from the Astrophysical Journal.“ (Kurtz und Henneken, 2007)

Zur Beurteilung von OA-Dokumenten eignet sich ebenfalls der Ansatz von (Kling und McKim, 1999). Die Autoren entwickeln in ihrem analytischen Ansatz die drei Dimensionen Publizität³¹, Zugänglichkeit³² und Vertrauenswürdigkeit³³, die helfen, elektronische Publikationen gegenüber herkömmlichen Publikationen zu bewerten.

„When a scholarly document is effectively published within a scholarly community, it seems to satisfy three criteria: publicity, access, and trustworthiness.“ (S. 897)

Nach einer eingehenden Analyse unterschiedlicher elektronischer Publikationsmodelle (z. B. Veröffentlichung von Publikationen in einem E-Journal vs. auf einem Dokumenten-Server vs. auf einer persönlichen Webseite) schlussfolgern die Autoren, dass die Publizität einer wissenschaftlichen Publikation nicht automatisch zunimmt, wenn das Dokument im Internet frei verfügbar vorliegt. Folglich sehen die Autoren auch keinen klaren Zusammenhang zwischen effektiver wissenschaftlicher Kommunikation und Web Posting-Strategien.

„Because few of the activities in the pre-publishing continuum promote strong publicity, trustworthiness, and access, they pose little threat to traditional journals“. (S. 903)

2.6.2 Integration von elektronischen Publikationen

Da sich der Anteil der rein elektronisch verfügbaren Publikationen durch OA und Self-archiving im Umfeld der Digitalen Bibliotheken stetig vergrößert und deren Akzeptanz steigt, stellt sich immer häufiger die Frage nach der Integrität und den Integrationsmöglichkeiten dieser neuen Publikationsformen. Wann kann ein elektronisches Angebot beispielsweise als wissenschaftlich integer bezeichnet werden? Unter welchen Bedingungen können derartige Publikationen in eine Digitale Bibliothek integriert werden? Kann die Integration erst nach vorheriger Integritätsprüfung erfolgen? Der pragmatische Vorschlag, dass eine Publikation - gleich welchen Typs - lediglich ein Peer-Reviewing

31 „The document is announced to scholars so that primary audiences and secondary audiences may learn of its existence. Publicity represents a continuum of activities from subscription, report lists, abstract databases, advertising and special issues, and citation.“ S. 897

32 „Readers must be able to access the document independent of the author, and, in a stable manner, over time. Accessibility is typically assured by institutional stewardship as practiced by libraries, publishing houses, clearinghouses, and is supported by stable identifiers, such as ISBN and ISSN.“ S. 897

33 „The document has been vetted through some social processes that assure readers that they can place a high level of trust in the content of the document based on community-specific norms. Trustworthiness is typically marked by peer review, publishing house/journal quality, and sponsorship.“ S. 897

durchlaufen haben muss, greift informationswissenschaftlich gesehen zu kurz. Es ist deshalb notwendig, operable Modelle und Lösungen zu entwickeln, die die Integration und Bündelung der elektronischen Publikationen befördern (siehe dazu als Beispiel die DRIVER-Infrastruktur³⁴ bei (Feijen et al., 2007)).

Die Auswirkung und die künftige Bedeutung von elektronischen Publikationen inkl. OA auf Suchsysteme wie vascoda oder sowiport können heute aufgrund der Zusammensetzung der Kollektionen nur sehr schlecht beurteilt werden. Gerade in den geistes- und sozialwissenschaftlichen Fächern sind die Möglichkeiten und Vorteile, Forschungsergebnisse OA oder elektronisch zu publizieren, wenig bekannt. Des Weiteren weisen die meisten heutigen Suchsysteme kaum frei verfügbare elektronische Publikationen nach. Für heutige Digitale Bibliotheken und Suchsysteme wird es insbesondere darum gehen, vertrauenswürdige und qualitativ hochstehende elektronische Angebote zu identifizieren und sie fachlich koordiniert in die bestehenden Systeme zu integrieren. Dies kann mit heutiger Technik technisch relativ leicht über standardisierte Pull-Mechanismen und Schnittstellen wie OAI-PMH³⁵ oder SRU³⁶ erfolgen. Rein automatische Ansätze ohne intellektuelle Qualitätsprüfung sind heute noch fehleranfällig, wie an den Informationsdiensten des Suchmaschinenanbieters Google (z. B. Google News oder auch Google Scholar, siehe nächster Abschnitt) gut gezeigt werden kann.

Wie (Kling und McKim, 1999) zeigen, fehlt es elektronischen Publikationen häufig noch an Publizität und Vertrauenswürdigkeit; beides kann z. B. durch kontrollierte Integration und geeignete Modellbildung erreicht werden (siehe qualitätsorientierter Ansatz in sowiport). Dies setzt aber notwendigerweise voraus, dass die neuen Systeme von den Wissenschaftlern und der interessierten (Fach-)Öffentlichkeit auch intensiv genutzt werden. Mehrwertdienste wie die in dieser Arbeit vorgestellten können dazu beitragen, die Attraktivität der künftigen Digitalen Bibliotheken durch mehr Nutzerfreundlichkeit zu erhöhen.

Nachdem in den vorherigen Abschnitten die typischen Informationsangebote und Informationsanbieter wissenschaftlicher Fachinformation vorgestellt wurden, ist der letzte Abschnitt dieses Kapitels der Untersuchung der fachübergreifenden Suchmaschine Google Scholar gewidmet. Google Scholar kann als ein Beispiel eines Allgemeinzugangs im Sinne von Abschnitt 2.2 angesehen werden.

2.7 Google Scholar als akademische Suchmaschine

Die Google Scholar (GS)-Suchmaschine wird in dieser Arbeit empirisch untersucht, weil sich das GS-Modell gut eignet, um die Vor- und Nachteile der rein automatischen und fachübergreifenden Dokumentenerschließung im Vergleich mit den traditionellen qualitativ-intellektuellen Verfahren zu diskutieren.

34 <http://www.driver-repository.eu/>

35 <http://www.openarchives.org/OAI/openarchivesprotocol.html>

36 <http://www.loc.gov/standards/sru/>

Der folgende Abschnitt beschreibt das GS-Modell und präsentiert die Ergebnisse der empirischen Studie, die 2006 im Rahmen des Projekts KoMoHe durchgeführt wurde (siehe dazu (Mayr und Walter, 2007)).

Der Start des Suchdienstes GS hat insbesondere wegen der Nähe zu den aktuell viel diskutierten Themen Open Access (vgl. Abschnitt 2.6.1) und Invisible Web (Bergman, 2001, Lewandowski und Mayr, 2006) für Aufsehen im Bereich der Fachinformation gesorgt. Google war 2004 der erste der drei großen kommerziellen Suchmaschinenanbieter, der sich mit seinem Dienst GS auf das Marktsegment „kommerzielle Fachinformation“ fokussiert. Google tritt damit in einen Markt ein, der seit den 1970er Jahren in der Hand von kommerziell orientierten Datenbankproduzenten ist, die für die Recherche in qualitativ erschlossenen Nachweisen von Forschungsliteratur Gebühren erheben (siehe dazu (Brophy und Bawden, 2005)).

Die Besonderheit von Google Scholar liegt neben der zugrunde liegenden Technologie sicherlich in seiner Bemühung, nur wissenschaftliche und qualitätsgeprüfte Dokumente zu durchsuchen. Die Beschränkung auf nachweislich wissenschaftliche Dokumente konnte bislang von keiner Internetsuchmaschine konsequent umgesetzt werden. Dieses Ziel versucht Google über Kooperationen mit einer größeren Zahl von Fachverlagen (z. B. Blackwell, Nature Publishing Group, Springer-Verlag usw.) und Fachgesellschaften (z. B. Association for Computing Machinery, Institute of Electrical and Electronics Engineers, Institute of Physics usw.) zu erreichen.

„Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations. Google Scholar helps you identify the most relevant research across the world of scholarly research.“ (Google, 2008)³⁷

Google Scholar selbst ist ein kostenfreier Service, der die gewohnte Google-Suche bereitstellt und über einen zentralen Index die Inhalte auf den unterschiedlichen Verlagsservern erschließt. Der Dienst ermöglicht es, auf Inhalte, die auf Verlagsservern gespeichert sind, direkt zuzugreifen. Im Idealfall kann ein Nutzer, der Zugriffsrechte auf einem Verlagssystem hat (z. B. durch institutionelle Subskriptionen), direkt auf den Volltextartikel zugreifen, den er über GS lokalisiert hat. Damit wird die Suche in den Fachdatenbanken der Datenbankanbieter, die die Fachliteratur ebenfalls erschließen, im Prinzip obsolet. Zusätzlich bietet der GS-Dienst die Möglichkeit, auf Volltexte im „freien“ Internet (insbesondere den Open Access- und Self-archiving-Bereich) zuzugreifen. Dies ist insbesondere für Personenkreise interessant, die über keine institutionellen Subskriptionen bzw. Mittel verfügen, um die Volltexte bei den Verlagen zu erwerben (siehe Abbildung 2-5).

Für Nutzer sind neben dem direkten Volltextzugang aber unter Umständen die von Google implementierten Mehrwertdienste und darauf aufbauend das Dokumentenranking interessant. Google Scholars Relevanzranking basiert laut eigenen Angaben auf

37 <http://scholar.google.de/intl/en/scholar/about.html>

unterschiedlichen Kriterien (siehe Zitat unten). Insbesondere die automatische Zitationsextraktion und -analyse, auch Autonomous Citation Indexing (ACI) genannt (Lawrence et al., 1999), kann für den Nutzer Hilfestellung bei der Informationssuche und -beschaffung bringen. Hochzitierte Arbeiten - Google nennt diese Arbeiten „key papers“ - werden nach diesem Verfahren oben in die Ergebnisliste gerankt und sind für Recherchierende damit gut sichtbar. Die Zitationen eines Treffers können angezeigt werden und damit kann ein Browsing durch die zitierenden Arbeiten angeboten werden. Das automatische Verfahren ACI setzt allerdings voraus, dass die Literaturangaben der analysierten Dokumente zur Verfügung stehen, was bei den Volltexten per se gegeben ist. Google Scholar kann damit über die Referenzen analysierter Dokumente hinaus auch Literaturquellen nachweisen, die nicht auf den indexierten Webservern liegen.

„Just as with Google Web Search, Google Scholar orders your search results by how relevant they are to your query, so the most useful references should appear at the top of the page. This relevance ranking takes into account the full text of each article as well as the article's author, the publication in which the article appeared and how often it has been cited in scholarly literature. Google Scholar also automatically analyzes and extracts citations and presents them as separate results, even if the documents they refer to are not online. This means your search results may include citations of older works and seminal articles that appear only in books or other offline publications.“ (Google, 2008)

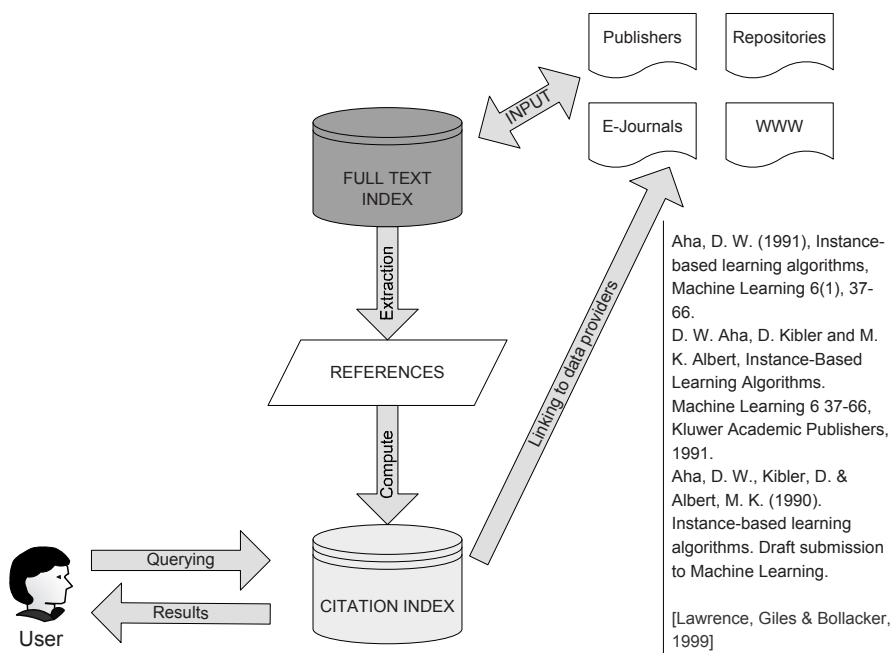


Abbildung 2-5: Google Scholar-Ansatz. Im rechten Bereich der Abbildung wird ein typisches Beispiel für unterschiedliche bibliographische Ansetzungen eines Aufsatzes präsentiert.

Weiterhin ist an Google Scholar interessant, dass diese Suchmaschine fachübergreifend konzipiert ist. Im Gegensatz zu Fachdatenbanken oder auch den Spezialsuchmaschinen wie z. B. dem CiteSeer-System, das frei verfügbare wissenschaftliche Informatikliteratur indexiert, oder RePEc für Arbeitspapiere in Economics, wäre mit dem Google Scholar-Ansatz eine umfassende Wissenschaftssuchmaschine für alle Disziplinen denkbar (vgl. Allgemeinzugang in Abschnitt 2.2). Abbildung 2-5 visualisiert den GS-Ansatz.

Da Google über die Reichweite, Aktualität und Abdeckung seines GS-Dienstes keine Informationen bereithält, sollte mit einer empirischen Studie untersucht werden, wie tief Google Scholar sich in das wissenschaftliche Web vorgearbeitet hat.

Ergebnisse der GS-Studie

Im Zeitraum August 2006 wurde der Umfang des Services anhand der Abdeckung unterschiedlicher Zeitschriftenlisten gemessen. Weiterhin wurde untersucht, welche Typen von Nachweisen und welche Webserver sich in den analysierten Trefferdaten befinden (vgl. (Mayr und Walter, 2007)).

Die Studie sollte Aussagen zu folgenden Fragen ermöglichen:

- Wie vollständig deckt Google Scholar die unterschiedlichen wissenschaftlichen Zeitschriften ab?
Die Studie testet über die Abfrage von unterschiedlichen Zeitschriftenlisten, ob Google die Zeitschriften indexiert hat und Artikel aus diesen Zeitschriften nachweisen kann. Die Zeitschriftenlisten kommen aus sehr unterschiedlichen Bereichen: internationale Peer-reviewed Zeitschriften des Web of Science³⁸ (überwiegend Science, Technology & Medicine), Open Access-Zeitschriften (DOAJ³⁹) und Zeitschriften der deutschsprachigen Sozialwissenschaften (IZ⁴⁰).
- Welche Dokument- bzw. Treffertypen sind in Google Scholar enthalten?
Die analysierten Trefferdaten geben Hinweise auf die Zusammensetzung der GS-Treffertypen: Link, Zitationsnachweis und Volltextlink (siehe Tabelle 3).
- Von welchen Anbietern kommen die meisten Dokumente?
Die Studie soll deutlich machen, wer die größten Datenlieferanten für den Suchdienst sind und welche wissenschaftlichen Informationsquellen aktuell im Index möglicherweise unterrepräsentiert sind. Die Verteilung der Webserver bzw. Anbieter ist interessant, weil sich daraus schließen lässt, ob Google Scholar eher kostenpflichtige Dokumente oder frei zugängliche erschließt.

Schwierigkeiten bei der Analyse traten bei der Überprüfung der Zeitschriftentitel auf. Beispielsweise gibt Google Scholar bei der Suche nach Artikeln der Zeitschrift „Applied Intelligence“ auch Treffer der Zeitschrift „Applied Artificial Intelligence“ aus, da exakte Phrasensuche zum Zeitpunkt der Untersuchung nicht möglich war. Ein weiteres Problem

38 <http://scientific.thomson.com/products/wos/>

39 <http://www.doaj.org/>

40 <http://www.gesis.org/dienstleistungen/fachinformationen/datenbanken-informationssysteme/literaturdatenbank-solis/>

stellt die uneinheitliche Darstellung der Titel dar, die auf die automatische Zitationsextraktion zurückzuführen ist: Zum Beispiel werden Artikel der Zeitschrift „Analyse und Kritik“ auch unter dem Titel „Analyse and Kritik“ oder „Analyse & Kritik“ aufgeführt.

Die aus den Trefferlisten extrahierten Daten wurden über einfache Auszählungen aggregiert. Die Treffer, die eindeutig einer Zeitschrift zugeordnet werden konnten, wurden den drei unterschiedlichen Treffertypen zugewiesen und ausgezählt (siehe Tabelle 2). Für jeden Treffer, der einer Zeitschrift zugeordnet werden konnte, wurden anschließend alle Domains (Webserver) extrahiert und die Häufigkeit der einzelnen Webserver pro Zeitschriftenliste bestimmt.

Nachfolgend finden sich einzelne Ergebnisse der Untersuchung.

Tabelle 2 zeigt die Anzahl der analysierten und identifizierten Zeitschriftentitel der unterschiedlichen Zeitschriftenlisten. Von den 317 Zeitschriften der deutschsprachigen sozialwissenschaftlichen Zeitschriften (GESIS) konnten beispielsweise 222 Titel (ca. 70 % der gesamten Liste) eindeutig identifiziert werden (siehe „Titel gefunden“). Die verbleibenden 30 % der Liste konnten nicht identifiziert werden oder haben keine Treffer in Google Scholar generiert. Die Zeitschriften von Thomson Scientific (Arts & Humanities Citation Index = AH, Science Citation Index = SCI, Social Science Citation Index = SSCI), die hauptsächlich englischsprachige Zeitschriften abdecken, zeigen die besten Abdeckungsquoten mit über 80 % identifizierten Titeln. Die Liste der Open-Access-Liste (DOAJ) liegt in etwa im Bereich der GESIS-Liste.

Tabelle 2: Anteil der identifizierten Zeitschriftentitel in den Google Scholar-Daten. Es wurden jeweils 100 Dokumente pro Zeitschriftenartikel analysiert.

| Liste | Titel | Titel gefunden (in %) |
|-------|-------|-----------------------|
| AH | 1.149 | 925 (80,50) |
| DOAJ | 2.346 | 1.593 (67,90) |
| GESIS | 317 | 222 (70,03) |
| SCI | 3.780 | 3.244 (85,82) |
| SSCI | 1.917 | 1.689 (88,11) |

Im Anschluss wurden die rund 621.000 Google Scholar-Treffer bzgl. Treffertypen untersucht (siehe Tabelle 3 mit den Anteilen je Zeitschriftenliste). Die Google Scholar-Treffer lassen sich in drei unterschiedliche Typen kategorisieren (Link, Citation und Volltext). Die Verteilung der Treffertypen hängt deutlich mit dem zuvor dargestellten Ergebnis zusammen. Der hohe Anteil der identifizierten Zeitschriften geht insbesondere auf einen hohen Anteil beim Treffertyp „Citation“ zurück. Die Treffertypen „Link“ und „Volltext“ führen direkt in die elektronischen Systeme der Anbieter oder eben zu den Volltexten der Artikel. „Citations“ können von Google nicht aufgelöst und referenziert werden und werden daher wie in den analysierten Dokumenten angegeben.

Tabelle 3: Anteil der Treffertypen in den Google Scholar-Daten

| Liste | Link in % | Citation in % | Volltext in % |
|-------|-----------|---------------|---------------|
| AH | 41,78 | 50,73 | 7,49 |
| DOAJ | 48,29 | 29,61 | 22,11 |
| GESIS | 10,42 | 83,11 | 6,48 |
| SCI | 61,35 | 16,72 | 21,94 |
| SSCI | 49,38 | 32,84 | 17,78 |

Fazit der GS-Studie

Der Anteil der deutschsprachigen wissenschaftlichen Zeitschriften (vgl. aktuelle Untersuchung Mayr & Umstätter, 2008) in Google Scholar, getestet anhand der sozialwissenschaftlich ausgerichteten Zeitschriftenliste der GESIS (83,11 % Citations), ist aller Wahrscheinlichkeit nach eher gering und unvollständig. Die Studie zeigt zwar, dass ein Großteil der Zeitschriften der analysierten Zeitschriftenlisten in den Google Scholar-Daten identifiziert werden kann, eine weitere Analyse der Treffertypen relativiert diese Ergebnisse aufgrund des hohen Anteils an extrahierten Referenzen (Treffertyp „Citation“) wieder.

Die Analyse der Webserver zeigt, dass vorrangig die Fachangebote der internationalen kommerziellen Wissenschaftsverlage wie z. B. Springer, Ingenta, Wiley usw. (allerdings nicht vollständig) indexiert wurden. Unsere Ergebnisse verdeutlichen, dass umfangreiche elektronisch frei zugängliche Bestände, insbesondere aus dem Open Access- (siehe DOAJ-Liste) und Self-archiving-Bereich im Untersuchungszeitraum zu wenig berücksichtigt wurden. Eine Nachfolgeuntersuchung der OA-Zeitschriften der DOAJ-Liste zeigt, dass die elektronischen Zeitschriften ohne Zugangsbeschränkung inzwischen von GS deutlich besser erfasst und indexiert werden als zum Zeitpunkt der GS-Studie im Jahr 2006.

Unsere Tests bestätigen, dass Google Scholar in vielen Dokumentkollektionen keine tagesaktuellen Daten präsentieren kann und die Trefferdaten aufgrund der Implementation der automatischen Zitationsextraktion z. T. unvollständig, fehlerhaft und häufig redundant aufgelistet werden (vgl. (Jacsó, 2005, Jacsó, 2008)).

Wie der bekannte Suchdienst Google Web Search bietet auch Google Scholar die gewohnt schnelle Suche und eine einfach zu bedienende Benutzeroberfläche. Pluspunkte sind, dass die Recherche kostenfrei ist und dass im Volltext fachübergreifender Bestände gesucht werden kann, was viele vergleichbare Systeme nicht ermöglichen. Der Ansatz von Google Scholar bietet für Literatursuchende einige Mehrwerte, wie z. B. die automatische Zitationsanalyse und das darauf aufbauende Ranking und Browsing sowie in vielen Fällen den direkten Volltextzugriff. Die Evaluation von Zitationszahlen oder webometrischen Untersuchungen (Thelwall et al., 2005) auf Basis der Google Scholar-Daten (vgl. (Bar-Ilan, 2006, Belew, 2005, Jacsó, 2004, Kousha und Thelwall, 2007)) wäre

aufgrund der kostenfreien Nutzung des Services u. U. fruchtbar, allerdings aufgrund der Vagheit in den Daten mit großer Vorsicht zu betrachten.

“Citation counts aggregated by Google Scholar may work in some fields that are covered and indexed quite well, but in other fields which are perhaps more represented by the freely accessible web, these counts can be very inflated. This can mislead researchers in citation analyses based solely on Google Scholar.” (Mayr und Walter, 2007)

Im Vergleich zu Fachdatenbanken mit ihren hohen Anforderungen an die Dokumentenqualität (z. B. nur peer-reviewed papers in WoS), Aktualität sowie der Fokussierung auf Precision und Recall bietet Google Scholar momentan nicht die Transparenz und Vollständigkeit, die viele Nutzer von einem wissenschaftlichen Informationsangebot erwarten. Als Ergänzung der Recherche in Fachdatenbanken - v. a. durch die Abdeckung einer Reihe von Open Access-Zeitschriften - kann Google Scholar aber durchaus nützlich sein.

Nachdem in diesem Kapitel die Breite, Vielfalt und unterschiedlichen Modelle der Informationsangebote und Informationsanbieter beschrieben wurden, soll im folgenden Kapitel konkreter auf die Bemühungen beim Zusammenführen von heterogenen Informationsangeboten eingegangen werden.

3 Heterogenität und Information Retrieval

Die Abschnitte dieses Kapitels konzentrieren sich auf die Darstellung der Ergebnisse des Projekts KoMoHe. Zuvor werden die grundlegenden Konzepte der Heterogenitätsbehandlung eingefügt und der Information Retrieval Test (siehe Kapitel 5) motiviert.

3.1 Schalenmodell

Für die Modellbildung in dieser Arbeit stellt insbesondere das Schalenmodell konzeptuell weiterführende Überlegungen an. Das von J. Krause (Krause, 1996) vorgeschlagene Modell eines Dokumentenraums entwickelt eine ganzheitliche Lösung, die hilft, die Informationsangebote, die bzgl. der Erschließung heterogen sind, strukturiert und reguliert anzuordnen (siehe Niveaus der Dokumentenrelevanz und Erschließung, Trennbarkeit der Schalen in Abbildung 3-2). Überarbeitete und erweiterte Fassungen des Schalenmodells finden sich in Krause (Krause, 2006, Krause, 2008). In diesen Papieren wird beispielsweise die parallele Entwicklung und Anschlussfähigkeit des Schalenmodells inkl. der Heterogenitätskomponenten (siehe Abschnitt 3.2.1) und der aktuellen Bemühungen beim Einsatz von Ontologien im Semantic Web hervorgehoben. Beide Ansätze fokussieren auf die qualitative Zusammenführung von Informationsbeständen auf Basis von kontrollierten Vokabularen. Das Schalenmodell setzt dazu vor allem auf Thesauri und Klassifikationen, während die Semantic Web Initiative deutlich formale deduzierbare Ontologien zugrundelegt.

„.. there is no fundamental contradiction between the bilateral transfer components ... and the ontological approach. ... the semantic knowledge of the thesauri and the classifications with the help of bilateral heterogeneous components can be used without blocking the way for subareas in which more in depth and logically more precise – but also more complex – ontological approaches are required.“ (Krause, 2006)

„Beide Ansätze, der des Semantic Web ... und der der semantischen Heterogenitätsbehandlung im Rahmen des Schalenmodells, teilen Grundprämissen, die sich gegen die Schwächen der Inhaltserschließungsverfahren der generellen Web-Suchmaschinen richten. Der Unterschied zwischen den Ansätzen, die parallel entstanden, liegt (bzw. lag bis vor kurzem) vor allem in der Art der als notwendig erachteten bzw. vorteilhaftesten ... semantischen Fundierung.“ ((Krause, 2008) S. 199)

Das ursprüngliche Schalenmodell geht ähnlich wie W. Umstätter (Umstätter, 2005) davon aus, dass künftig wissenschaftliche Informationsangebote aus dem Bibliotheks- und Dokumentationsbereich, die i. d. R. für Internetsuchmaschinen nicht erreichbar sind (siehe Invisible Web bei (Bergman, 2001)), dem Nutzer gemeinsam in der „Digitalen Bibliothek“ angeboten werden müssen.

Folgende essentielle Bedingungen müssen Informationsangebote erfüllen, um in das Schalenmodell integriert werden zu können: fachwissenschaftliche Relevanz der Quellen und koordinierte und kontrollierte Integration der deregulierten Angebote. Wünschenswert ist weiterhin ein konsistenter und tief erschlossener Kernbereich (Schale 1), der sich von weiteren Schalen bzgl. der Qualität der Inhaltserschließung unterscheidet. Es

bleibt schließlich den Informationsanbietern bzw. der Koordinierungsinstanz überlassen, wie viele Schalen angesetzt und welche Kriterien für die Definition der Schalen genutzt werden. So kommt Krause aus Sicht eines Fachdatenbankanbieters zu folgenden Schaleneinteilungen für die Sozialwissenschaften:

„Die innerste Schale enthält den Kern der relevanten Literatur. Er wird möglichst tief und qualitativ hochwertig erschlossen. Die Qualitätskontrolle liegt in der Hand der koordinierenden Informationsservicestelle. Die zweite Schale lockert die Relevanzbedingungen und parallel dazu die Anforderungen an die Qualität der Inhaltserschließung. Schale 3 könnte alle Dokumente enthalten, deren Relevanz gegenüber den Schalen 1 und 2 niedriger ist und die nach anderen Normen erschlossen sind (z. B. anderer Thesaurus). Schale 4 enthielte die Ansetzung der Bibliotheken. Neben den gebundenen Deskriptoren (Beispiel Autor) steht für die Inhaltserschließung nur der Titel zur Verfügung, der automatisch indexiert wird.“ (Krause, 1996)

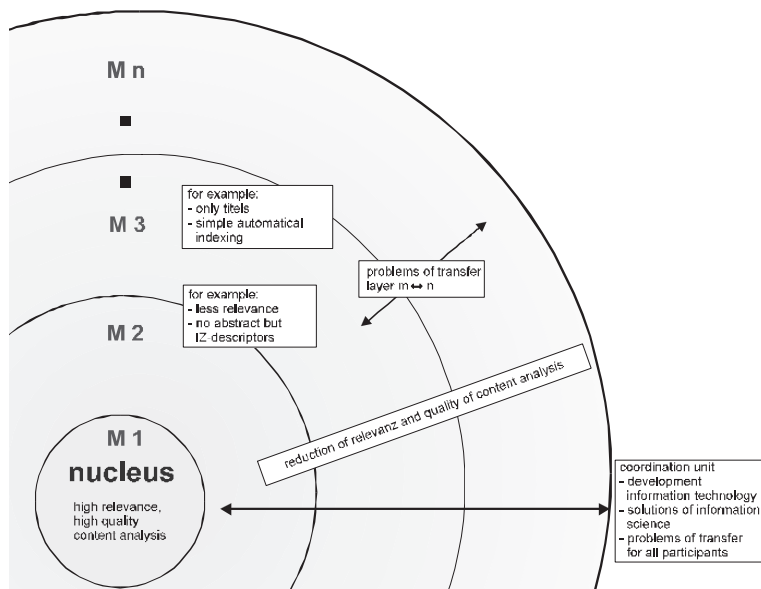


Abbildung 3-1: Schalenmodell (aus Krause, 1996)

Abbildung 3-1 präsentiert eine typische Benutzersicht auf die heutige Informationslandschaft, die das Schalenmodell zur Grundlage der Organisation der einzelnen, unterschiedlich erschlossenen Informationsangebote macht. Charakteristisch ist die dezentrale Struktur (polyzentrisch, vgl. (Krause, 2003)) mit einer Vielzahl von infrage kommenden Informationsanbietern und Dokumenttypen (siehe Abschnitt 2.5). Im oberen Bereich der Abbildung finden sich die klassischen Fachinformationsanbieter wie zum Beispiel Informationszentren, Verlage, Bibliotheken und die Wissenschaftler selbst, die alle relativ unabhängig voneinander Dokumente produzieren und in Form von Datenbanken oder

anderweitig in mehr oder weniger aggregierter Form zur Verfügung stellen. Die dabei angewendeten Erschließungsverfahren variieren zwischen den einzelnen Angeboten, aber auch zwischen den einzelnen Disziplinen bzw. Domänen.

Vorteil des Schalenmodells ist neben der Selbstbestimmung der Teilbereiche, dass die Schalendefinitionen alle wichtigen Erschließungsmodelle einbeziehen. Neben den beiden intellektuellen Erschließungstraditionen: generelle Sacherschließung der Bibliotheken durch Normdaten und Klassifikationen und fachspezifische Indexierung der Information und Dokumentation (IuD) durch Thesauri (siehe dazu Einsatzmöglichkeiten bei (Blocks, 2004, Tudhope et al., 2006)) findet auch die automatische Erschließung, z. B. durch Suchmaschinen, Eingang in das oben vorgeschlagene Modell. Damit ist das Modell selbst für „verschmutzte“, rein automatisch (algorithmisch) erschlossene Dokumente der äußeren Schalen offen. Das Modell ist zudem in der Lage, neue Erschließungsmodelle wie z. B. Tagging (Voß, 2007) in das Gesamtkonzept einzugliedern. Im Schalenmodell werden die Schalenübergänge idealerweise durch die Komponenten zur Heterogenitätsbehandlung mittels einer Transfer- und Koordinationsschicht überwunden (siehe Abbildung 3-2).

Alternativ-Modelle wie die Standardisierung der kontrollierten Vokabulare in Form eines Metathesaurus (siehe dazu (Fenske, 2006)) – für den Bereich der Medizin z. B. das Unified Medical Language System (UMLS) – scheiden als Modell für interdisziplinäre Systeme im Grunde aus. Es wird nicht angenommen, dass es finanziell, organisatorisch und politisch möglich und sinnvoll ist, einen gemeinsamen Thesaurus für mehrere Fächer durchzusetzen. Gleiches gilt für den breiten Einsatz von Ontologiesprachen des Semantic Web.

Die Grundidee des Schalenmodells lässt sich dahingehend untersuchen, ob die Forderung zur Erstellung „konsistenter und tief erschlossener“ Kernbereiche – in der Regel eine oder mehrere integrierte Fachdatenbanken aus einer Disziplin – durch operationalisierbare Kriterien evaluierbar wird. Fachwissenschaftliche Relevanz, Konsistenz und tiefe Erschließung sind zunächst rein konzeptuell, können aber durchaus in Teilbereichen operationalisiert und untersucht werden. Bradford's Law of Scattering (BLS) bzw. dessen Anwendung im Verfahren Bradfordizing (siehe Kapitel 4) ist eine Möglichkeit, um die oben genannten Anforderungen des Schalenmodells zu untersuchen. Dies soll im Evaluationskapitel dieser Arbeit geschehen (siehe Kapitel 5.2).

Diese Überlegung basiert auf der Analogie von Schalenmodell (Kernbereiche der Forschungsliteratur) und den Kernbereichen von Bradford's Law of Scattering. Die bei Bibliotheks- und Informationswissenschaftlern wohlbekannte informetrische Gesetzmäßigkeit des BLS beschreibt ebenfalls Kernzonen, allerdings anhand von Häufigkeitsverteilungen in Fachzeitschriften. Da das Schalenmodell in seiner inhaltlichen Auslegung bereits auf die tiefer erschlossene Kernzone mit Dokumenten aus Fachdatenbanken fokussiert, lässt sich die Erweiterung um weitere Kernbereiche (Bradford Zone 1 bzw. Nukleus) innerhalb der Kernzone des Schalenmodells gut motivieren (siehe dazu Abschnitt 3.1).

Das Schalenmodell (siehe insbesondere die Erweiterungen in (Krause, 2006, Krause, 2008)) ist eine wichtige konzeptuelle Basis für das Projekt KoMoHe, das im folgenden Abschnitt beschrieben wird.

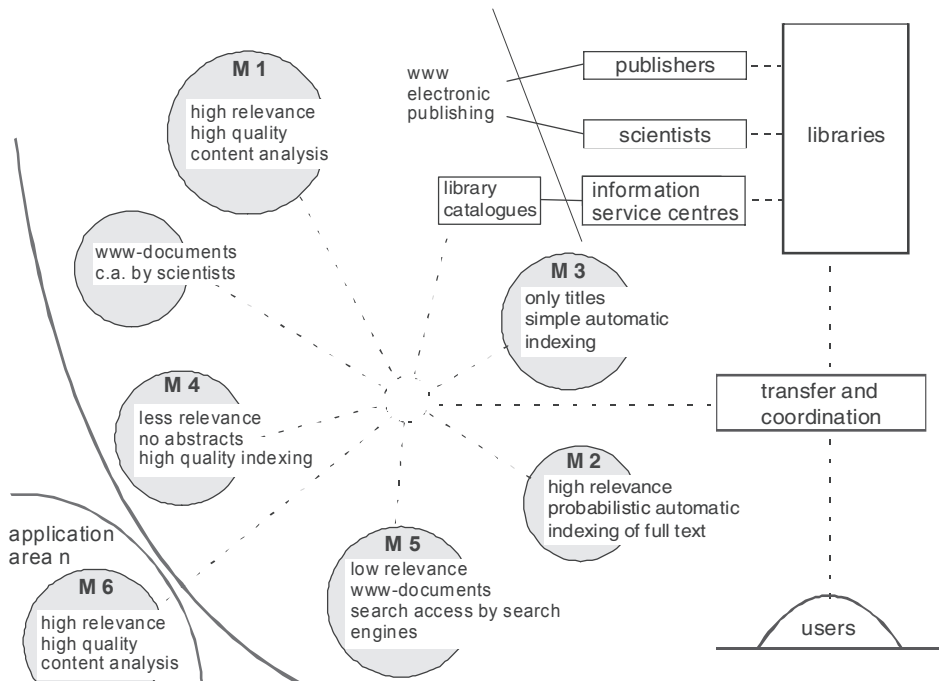


Abbildung 3-2: Polyzentrischer Dokumentenraum im Schalenmodell (aus Krause, 2003)

3.2 Projekt KoMoHe

Die Behandlung und Modellierung der Heterogenität wissenschaftlicher Dokumentensammlungen stand im Fokus des Teilprojekts „Kompetenzzentrum Modellbildung und Heterogenitätsbehandlung“ (KoMoHe⁴¹) der GESIS⁴²-Abteilung „Informationelle Prozesse in den Sozialwissenschaften“ (vgl. dazu den Leitgedanken „Standardisierung von der

41 <http://www.gesis.org/forschung-lehre/programme-projekte/informationswissenschaften/projektuebersicht/komohe/>

42 GESIS (Abteilung „Fachinformation für Sozialwissenschaften“) in Bonn dokumentiert die Ergebnisse sozialwissenschaftlicher Forschung im deutschsprachigen Raum und macht sie im Rahmen elektronischer, gedruckter und persönlicher Informationsprodukte und Dienstleistungen Nutzern und Nutzerinnen aus Wissenschaft und Praxis zugänglich. GESIS in Bonn besteht aus zwei Abteilungen, die für Fachinformationsdienste und für die informationswissenschaftliche Forschung und Entwicklung (Abteilung „Informationelle Prozesse in den Sozialwissenschaften“) der gesamten GESIS zuständig sind. Die Informationsdienstleistungen basieren größtenteils auf den Datenbanken zu Forschungsprojekten (SOFIS) und Literatur (SOLIS), die als eigenständige Angebote über Hosts und in von GESIS und seinen Partnern entwickelten Fachinformationsangeboten im Internet verfügbar sind. Die Projektergebnisse ihrer informationswissenschaftlichen Forschung und Entwicklung stellt GESIS nicht-kommerziellen Einrichtungen grundsätzlich kostenlos zur Verfügung. Der Schwerpunkt der Forschungs- und Entwicklungsarbeiten liegt auf der intelligenten Integration heterogen strukturierter und

verbleibenden Heterogenität her denken“ in (Krause, 2003, Krause, 2004, Krause et al., 2003)).

Das KoMoHe-Projekt wurde vom BMBF im Zeitraum September 2004 bis August 2007 gefördert⁴³ und versucht, ein informationswissenschaftlich basiertes Rahmenkonzept bzw. Modell für Dokumentenräume auf Basis des Schalenmodells zu entwickeln. Die GESIS-Arbeitsgruppe bestand aus folgenden Personen: Jürgen Krause, Maximilian Stempfhuber, Vivien Petras, Anne-Kathrin Walter, Stefan Bärish und Philipp Mayr. Die in diesem Projekt entstandenen Modelle und Teilmodule sind prototypisch am Beispiel von *vascoda* entwickelt worden und sollen beschreiben, wie einzelne fachliche Informationsangebote zu einer interdisziplinären Informationsarchitektur zusammengefügt werden können (siehe dazu Abschnitte 3.1 und 4.4). Einzelne im Projekt erarbeitete Verfahren konnten zudem empirisch evaluiert werden. Die Evaluation des Verfahrens *Bradfordizing* (siehe dazu nächstes Kapitel 4) steht im Mittelpunkt dieser Arbeit.

Der zweite wichtige Bestandteil des Projekts bestand in einer Initiative zum Terminologie-Mapping zwischen kontrollierten Vokabularen (siehe folgender Abschnitt). Eine Aufgabe dieser Initiative war die Erstellung von Crosskonkordanzen (semantische Querverbindungen) zwischen kontrollierten Vokabularen innerhalb der Sozialwissenschaften und anderen Fachgebieten. Das Hauptziel des Projekts war die Erstellung, Implementation und Evaluation eines Terminologie-Netzes (Mayr und Petras, 2008) sowie die Ermöglichung semantischer Integration bei der Suche nach heterogenen Quellen in einem typischen interdisziplinären Suchsystem wie z. B. *vascoda* (vgl. (Chan und Zeng, 2006)).

Das Ziel der semantischen Integration ist es, unterschiedliche Informationssysteme über ihre inhaltlichen Metadaten zu verbinden und damit die verteilte Suche über mehrere Informationsangebote zu ermöglichen, ohne die erweiterten Möglichkeiten des Fachzugangs der individuellen Datenbanken zu beschränken. Durch den Überstieg zwischen verschiedenen Fachterminologien wird eine „semantische Kongruenz“ für gemeinsam recherchierbare Kollektionen erreicht. Terminologie-Mapping – die Abbildung von Begriffen und Phrasen eines kontrollierten Vokabulars auf Begriffe und Phrasen eines anderen kontrollierten Vokabulars – ermöglicht den reibungslosen Überstieg von einer Ein-Datenbank-Suche in ein verteiltes Suchszenario mit mehreren Datenbanken.

Der folgende Abschnitt erläutert die unterschiedlichen Verfahren der Heterogenitätsbehandlung, die im Projekt KoMoHe erprobt und untersucht wurden. Der Schwerpunkt der Darstellung liegt auf den intellektuellen Verfahren (Crosskonkordanzen), die auch im Fokus der Terminologie-Mapping-Initiative standen. Es werden die involvierten Vokabulare und Datenbanken sowie die Implementierung der erarbeiteten Crosskonkordanzen für die Suche beschrieben.

erschlossener Informationsbestände mit dem Ziel, den Nutzern und Nutzerinnen einen auf allen Ebenen einheitlichen Zugriff auf die von ihnen benötigten Fachinformationen zu bieten, gleich aus welcher Quelle sie stammen.

43 BMBF-Förderkennzeichen: 523-40001-01C5953

3.2.1 Heterogenitätsbehandlung

Eines der klassischen Probleme im Information Retrieval ist die Vagheit, die zwischen einer Anfrage eines Benutzers und den Termen, die ein Dokument beschreiben, besteht. Konkret bedeutet dies, dass ein Benutzer für eine beliebige Fragestellung häufig andere Terme verwendet als derjenige, der die Inhaltserschließung für entsprechende Dokumente vornimmt (vgl. language problem bei (Blair, 1990, Buckland, 1999, Buckland und Gey, 1999, Petras, 2006)). Die Vagheit zwischen Anfrage- und Dokumentebene wird bei Krause „Vagheit 1“ (V1) genannt (siehe Abbildung 3-3) und kann durch Verfahren zur Termerweiterung behandelt werden (Krause, 2003). Lösungsvorschläge für die Vagheitsbehandlung auf der Ebene V1 finden sich z. B. bei V. Petras (Petras, 2006). Als Instrument zum Vorschlagen besserer Suchterme reiht sich der Search Term Recommender (STR) in die Gruppe der Werkzeuge zur Suchverbesserung und Vermittlung zwischen Nutzer und Informationssystem ein (siehe auch (Plaunt und Norgard, 1998)). Durch eine Anfrageerweiterung mit Termen aus dem kontrollierten Wortschatz der inhaltlichen Beschreibungssysteme der Datenbanken reduziert der STR die Heterogenität der eingesetzten Terminologien (durch Angleichung der Suchsprache an die kontrollierten Vokabulare) und verbessert damit die Vagheitsrelation V1 in Abbildung 3-3. Gegenüber anderen Anfrageexpandierungstechniken oder einfachen Termvervollständigungsmodulen (z. B. Googles Vorschlags-Feature) hat der STR den entscheidenden Vorteil, dass Terme aus dem bei der Inhaltserschließung verwendeten kontrollierten Vokabular vorgeschlagen werden. Dies erhöht die Erfolgchancen, relevante Dokumente zu einem bestimmten Konzept zu finden (sind doch nur die Dokumente, die ein bestimmtes Konzept behandeln, mit dessen Schlagwort indexiert).

Handelt es sich bei den zu durchsuchenden Datenbeständen um homogen erschlossene Datenbanken, sind die Verfahren zur Behandlung von V1 ausreichend. Bei heterogen erschlossenen Dokumentenbeständen (vgl. Abschnitt 3.2.1) sind zusätzliche Verfahren anzuwenden. Durch den Einsatz von unterschiedlichen Thesauri und kontrollierten Vokabularen bei der Indexierung entsteht Vagheit bzw. Heterogenität bereits auf der inhaltlichen Beschreibungsebene der Dokumente. Beispielsweise wird ein Dokument bei der Sacherschließung mit Thesaurus A mit dem Deskriptor „Biologieunterricht“ beschrieben, während ein anderes kontrolliertes Vokabular B vorschreibt, in diesem Fall den Deskriptor „naturwissenschaftlicher Unterricht“ zu verwenden (siehe dazu Tabelle 6 und Zitat unten). Die Behandlung dieser Vagheit geschieht bilateral zwischen den einzelnen Dokumentbeständen, wie durch Abbildung 3-3 (**Vagheit 2 und 3 bzw. V2, V3**) verdeutlicht wird.

“Vocabulary problems are increased when dealing with foreign languages. (...) So „Drachenphotographie,“ in this specialized context, referred not to photography from a kite, but from an observation balloon. Even knowing that, the searcher would have been more effective if he had known that, in English, the term „Aerostat“ was used as a technical for observation balloons.” (Buckland, 1999)

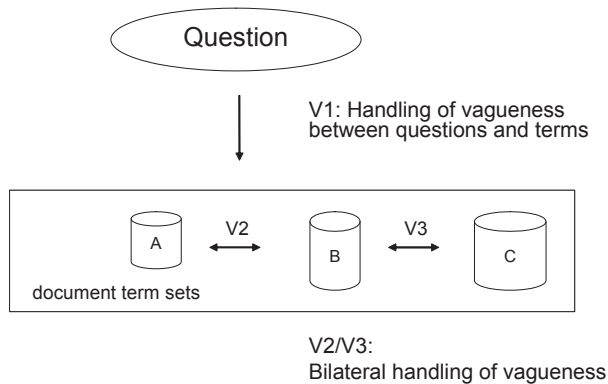


Abbildung 3-3: Vagheitsbehandlung im Information Retrieval (Zwei-Schritt-Modell aus (Hellweg et al., 2001))

Im oberen Beispiel (Abbildung 3-3) bedeutet das, dass zur Behandlung von V2 alle Terme aus Vokabular A auf Terme in Vokabular B abgebildet werden und umgekehrt alle Terme aus Vokabular B auf A (siehe dazu Abschnitt 3.2.2).

Semantische Interoperabilität kann auf unterschiedlichen Wegen erreicht werden. Ein guter Überblick über unterschiedliche Terminologie-Mapping-Methodologien und Projekte findet sich in (Zeng und Chan, 2004) und M. Doerr (Doerr, 2001, Doerr, 2004).

Bei GESIS sind unterschiedliche Ansätze erprobt worden, um semantische Heterogenität in Digitalen Bibliotheken zu behandeln (siehe dazu (Binder et al., 2002, Hellweg et al., 2001)).

- Crosskonkordanzen zwischen kontrollierten Vokabularen: Die verschiedenen Begriffssysteme werden im Kontext der Benutzer analysiert und es wird intellektuell versucht, die Konzeptualisierung der Begriffe miteinander in Verbindung zu bringen. Dieses Konzept darf nicht mit der Konstruktion von Metathesauri verwechselt werden. Beim Erstellen von Crosskonkordanzen wird nämlich kein Versuch unternommen, die existierenden Begriffswelten zu standardisieren. Crosskonkordanzen gewährleisten nur eine partielle Vereinigung der bestehenden Terminologiesysteme. Sie überbrücken den verbleibenden statischen Teil (Vokabulare) der Transferproblematik zwischen diesen Systemen. Diese Konkordanzen liefern meist Abbildungen (vgl. Tabelle 5, Tabelle 6, Tabelle 7) im Sinne von Synonym-, Ähnlichkeits- oder Hierarchie-Relationen, die auch als deduzierbare Regelrelationen eingesetzt werden können (Krause, 2008).
- Quantitativ-statistische Verfahren: Das Transferproblem kann generell als ein Vagheitsproblem zwischen zwei Inhaltsbeschreibungssprachen modelliert werden. Um die Vagheit beim Information Retrieval zwischen Termen, z. B. zwischen denen der Benutzeranfrage und denen der Dokumentkollektionen zu behandeln, sind verschiedene automatische Verfahren (probabilistische Prozeduren, Fuzzy-Verfahren und neuronale Netze) vorgeschlagen worden, die beim Termtransfer genutzt werden können (Hellweg et al., 2001, Lauser et al., 2008). Dasselbe individuelle Dokument kann gleichzeitig

von zwei Konzeptsystemen indexiert werden und somit können auf Basis der automatischen Verfahren unterschiedliche Systeme, die jeweils anders indexieren, miteinander in Verbindung gebracht werden (vgl. (Marx, 2005, Strötgen, 2007, Zhang, 2006)). Diese Verfahren benötigen allerdings Trainingsdaten. Beim multilingualen IR kann derselbe Text z. B. in zwei unterschiedlichen Sprachen sein.

- Deduktiv: Bei deduktiven Verfahren wird Textmaterial analysiert und aus den sich ergebenden Zusammenhängen werden mit Hilfe von logischen Schlussfolgerungen Relationen zwischen Termen abgeleitet.

Essentiell für alle Bemühungen des Terminologie-Mapping ist, dass verbleibende und unveränderbare Unterschiede zwischen verschiedenen Terminologien akzeptiert werden müssen. Aufgrund von Qualitäts- und Kostengründen ist keines der oben genannten Verfahren allein für die Transferproblematik zwischen heterogenen Kollektionen verantwortlich. Nach (Krause, 2003) sollten sich die Verfahren gegenseitig komplementieren und integriert sein.

Der Schwerpunkt der im Projekt KoMoHe erstellten Term-Transformationen liegt auf den intellektuell erstellten Crosskonkordanzen, daher wird darauf im Folgenden näher eingegangen.

Wenn semantische Heterogenitätsbehandlung (z. B. durch die Nutzung von intellektuellen Crosskonkordanzen) in einem verteilten Suchszenario implementiert ist, kann das System mit den verschiedenen Informationskollektionen mit dem Fachvokabular durchsucht werden, mit dem der Suchende vertraut ist. Crosskonkordanzen können damit die verteilte Suche auf unterschiedliche Arten unterstützen. Zuerst und vor allem soll die nahtlose Suche in Datenbanken mit unterschiedlichen kontrollierten Vokabularen ermöglicht werden. Zusätzlich können die Terminologie-Überstiege als Tool für die Vokabular-Erweiterung dienen, da sie ein Terminologie-Netz aus äquivalenten, weiteren und engeren sowie verwandten Term-Verbindungen repräsentieren (siehe dazu die Termbeispiele in Tabelle 4). Drittens kann das Vokabular-Netz mit seinen semantischen Abbildungen auch zur Expansion und Reformulierung von Anfragen genutzt werden.

Für den Fall, dass andere Informationskollektionen in die Digitale Bibliothek eingebunden sind, wird die Anfrage nicht nur in präzisen Suchstatements formuliert, sondern durch den Terminologie-Service automatisch in alle implementierten Terminologien übersetzt. Ein Recherchierender kann damit nahtlos zwischen verschiedenen Informationsressourcen wechseln, weil die semantische Übersetzung zwischen den unterschiedlichen Terminologien automatisch ausgeführt wird.

Für interdisziplinäre Informationssysteme (Beispiel vascoda) erhöht die semantische Integration z. B. durch Crosskonkordanzen nicht nur die Erfolgchancen verteilter Suchen über Kollektionen mit unterschiedlichen Erschließungsschemata, sondern eröffnet dem Suchenden einen Einblick in die anderen disziplinären Kontexte und die domänenspezifische Sprache, vorausgesetzt, die abgebildeten Vokabulare werden zur Verfügung gestellt.

Semantische Abbildungen spielen zusätzlich eine große Rolle beim Anbieten von Transfer-Methodologien zwischen fremdsprachigen Datenbanken (siehe z. B. Vorstudie in Kapitel 6). So wie ein Mapping zwischen kontrollierten Vokabularen unterschiedlicher Datenbanken oder Disziplinen erstellt werden kann, können Crosskonkordanzen

auch eine Übersetzung im klassischen Sinne bieten: zum Beispiel in Tabelle 4 von einer deutschsprachigen Terminologie zu einer englischsprachigen.

Tabelle 4 präsentiert zwei Ausgangsterme (linke Spalte) im deutschsprachigen Thesaurus Sozialwissenschaften (TheSoz) und intellektuell verbundene Deskriptoren (Endterme) in den abgebildeten Vokabularen (Endvokabularen). Die Relationstypen zwischen den Termen werden genauer erklärt in Tabelle 5.

Tabelle 4: Start- oder Ausgangsterm im TheSoz-Vokabular und eine Auswahl von Endtermen (semantische Mappings). Die Auflösungen der Kürzel der Endvokabulare finden sich in Tabelle 9.

| Startterme TheSoz | Relation | Endterme | Endvokabulare |
|--|----------|-----------------------------------|---|
| Weiterbildung engl.: „further education“ | = | Weiterbildung | Psyndex, STW, Info- data, SWD, BISp, DZI |
| | ^ | Berufsfortbildung | FES |
| | = | Further education | CSA-ASSIA |
| | = | Continuing education | CSA-PEI |
| | = | Adult Education | CSA-SA |
| | < | Education | CSA-WPSA |
| | = | Erwachsenenbildung | IBLK |
| Meinungsforschung engl.: „opinion research“ | 0 | | Psyndex |
| | ^ | Einstellungsforschung | IAB |
| | = | Opinion Polls | CSA-ASSIA |
| | = | Opinions + Research | CSA-SA |
| | < | Research | CSA-PEI |
| | = | Public Opinion Research | CSA-WPSA |
| | = | Public Opinion Polls | ELSST |
| | = | Meinungsumfrage/Meinungsforschung | IBLK |

In den letzten Jahren haben unterschiedliche Institutionen Bemühungen im Bereich der semantischen Integration von Informationssystemen unternommen. Das Projekt KoMoHe konnte insbesondere auf Erfahrungen in den Vorgängerprojekten ViBSoz, CARMEN und infoconnex der GESIS zurückgreifen.

Beispiele für das Mapping kontrollierter Vokabulare im internationalen Kontext sind:

- Mappings, die bei OCLC erstellt wurden (s. (Godby et al., 2004, Vizine-Goetz et al., 2004)): DDC (Dewey Decimal Classification) mit LCC (Library of Congress Classifica-

tion), Medical Subject Headings (MeSH) und NLM classification, sowie ERIC thesaurus mit LCSH (LC Subject Headings) und MeSH mit LCSH

- Mappings der Food and Agriculture Organization of the United Nations (FAO): AGRO-VOC zu Chinese Agricultural Thesaurus (s. (Liang und Sini, 2006))
- Mappings im MACS-Projekt⁴⁴ (Freyre und Naudi, 2001): Konkordanz zwischen Schlagwortnormdatei (Deutsch), RAMEAU (Französisch) und LCSH (Englisch)
- Mappings im Projekt Renardus⁴⁵ (Koch et al., 2001): DDC zu kontrollierten Vokabularen

In den USA hat OCLC das Terminology Services-Projekt⁴⁶ gestartet, um Web Services für Terminologie-Mappings zwischen unterschiedlichen kontrollierten Vokabularen wie z. B. DDC, LCC, LCSH oder MeSH anzubieten. In Europa hat das Delos2 Network of Excellence on Digital Libraries-Programm ein Arbeitspaket dem Problem von Wissensextraktion und semantischer Operabilität gewidmet (Patel et al., 2005). Ein anderer Bericht, beauftragt durch JISC, bietet einen Überblick über Terminologie Services mit Fokus auf Projekte in Großbritannien (Tudhope et al., 2006). Andere Projekte sind das CRISSCROSS⁴⁷-Projekt an der Deutschen Nationalbibliothek und der Fachhochschule Köln, das ein multilinguales thesaurus-basiertes Forschungsvokabular zwischen der Schlagwortnormdatei (SWD) und den Notationen der Dewey Dezimal Klassifikation (DDC) erstellt (Panzer, 2008). Die Agricultural Information Management Standards-Abteilung der FAO⁴⁸ ist in unterschiedliche Terminologie-Mapping-Initiativen involviert (z. B. (Lauser et al., 2008, Liang und Sini, 2006)). Das High-Level Thesaurus Project (HILT⁴⁹) an der University of Strathclyde ist ein weiteres Projekt mit langjähriger Erfahrung bei der Entwicklung von Terminologie-Mapping-Technologien (Macgregor et al., 2007).

3.2.2 Erstellung von Crosskonkordanzen

Wir definieren Crosskonkordanzen als intellektuell und manuell erstellte Verbindungen, die Äquivalenz, Hierarchie und Verwandtschaft zwischen Termen zweier kontrollierter Vokabulare über Relationen bestimmen. Typischerweise werden die Vokabulare bilateral verbunden, d. h. eine Konkordanz verbindet Terme eines Vokabulars A zu einem Vokabular B und eine weitere Konkordanz verbindet Terme von Vokabular B zurück zu A. Bilaterale Relationen sind dabei nicht notwendigerweise symmetrisch. Beispielsweise wird der Term ‚Computer‘ aus Vokabular A auf den Term ‚Information System‘ in Vokabular B abgebildet, aber der gleiche Term ‚Information System‘ in Vokabular B kann mit einem anderen Term, z. B. ‚Data base‘, in Vokabular A relationiert werden.

44 <https://macs.vub.ac.be/pub/>

45 <http://renardus.sub.uni-goettingen.de/>

46 <http://www.oclc.org/research/projects/termservices/>

47 <http://www.d-nb.de/eng/wir/projekte/crisscross.htm>

48 <http://www.fao.org/aims/>

49 <http://hilt.cdli.strath.ac.uk/>

Das Verfahren erlaubt die folgenden 1:1- oder 1:n-Relationen:

- Äquivalenz (=) bedeutet Identität, Synonymie oder Quasi-Synonymie
- Hierarchie (weitere Terme <; engere Terme >)
- Verwandtschaft (^) für ähnliche Terme
- Eine Ausnahme ist die Null (0)-Relation, die bedeutet, dass ein Term nicht auf einen anderen Term abgebildet werden kann (z. B. Relation Nummer 4 in Tabelle 5).

Zusätzlich muss jede Relation mit einem Relevanz-Rating (hoch, mittel und gering) versehen werden. Die Relevanz-Ratings sind ein sekundäres und relativ schwaches Instrument, um die Qualität der Relationen auszudrücken. Sie werden in unseren aktuellen Implementationen nicht ausgewertet.

Tabelle 5 präsentiert typische Crosskonkordanzen zwischen zwei Vokabularen A und B.

Tabelle 5: Crosskonkordanz-Beispiele

| Nr. | Vokabular A | Relation | Vokabular B | Beschreibung |
|-----|----------------------|----------------|----------------------|--|
| 1 | hacker | = | Hacking | Äquivalenz-Relation (=) |
| 2 | hacker | ^ ₊ | computers + crime | 2 Verwandtschafts-Relationen (^) über eine Termkombination (1-n) |
| 3 | hacker | ^ ₊ | Internet + security | |
| 4 | isdn device | 0 | | Null-Relation (0). Das Konzept kann nicht abgebildet werden. Der Term ist ev. zu spezifisch. |
| 5 | isdn | < | telecommunications | Engerer Term zu Oberbegriff (<) |
| 6 | documentation system | > | abstracting services | Weiterer Term zu Unterbegriff (>) |

Die Crosskonkordanzen im Projekt KoMoHe involvieren die gesamten oder umfangreiche Teile der Vokabulare. Bevor das Mapping der Terme beginnt, werden die Vokabulare bezüglich thematischen und syntaktischen Überlappungen untersucht. Die Crosskonkordanzen werden von Wissenschaftlern oder Terminologie-Experten innerhalb einer Domäne erstellt. Es ist essentiell für ein erfolgreiches Mapping, dass die Bedeutung und Semantik der Terme und der internen Relationen der beteiligten Vokabulare vollständig verstanden werden. Insbesondere das Verstehen und semantisch korrekte Relationieren der Terme unterscheidet die intellektuellen Verfahren von den rein automatischen Ansätzen (siehe dazu (Lauser et al., 2008)). Das korrekte Relationieren setzt syntaktische Prüfung von Wortstämmen und semantisches Verständnis bei der Suche nach Synonymen und anderen verwandten Begriffen voraus. Der Mapping-Prozess basiert auf einem Set von praktischen Regeln und Richtlinien (siehe beispielsweise (Patel et al., 2005)). Während des Mappings der Terme werden alle Intra-Thesaurusrelationen (Scope Notes eingeschlossen) verwendet. Recall und Precision der erstellten Relationen sollen in

den entsprechenden Datenbanken geprüft werden. Dies ist insbesondere für Kombinationen von Termen (1:n-Relationen) wichtig. 1:1-Termrelationen sollen bevorzugt werden. Wortgruppen und Relevanz-Ratings müssen konsistent umgesetzt werden. Zum Schluss wird die semantische Korrektheit der Crosskonkordanzen von Experten kontrolliert, zudem werden Stichproben empirisch auf Dokument-Recall und -Precision geprüft.

3.2.2.1 Einsatz von Crosskonkordanzen am Beispiel infoconnex

Das Angebot von infoconnex ist ein interdisziplinärer Informationsverbund für die Fächer Pädagogik, Sozialwissenschaften und Psychologie, der die integrierte Recherche in den zentralen Fachdatenbanken der drei Disziplinen ermöglicht. Für die integrierte Recherche über die drei Datenbanken (FIS Bildung, SOLIS und Psyndex) werden Crosskonkordanzen eingesetzt.

Tabelle 6 zeigt ausgewählte Deskriptoren des Thesaurus Sozialwissenschaften (The-Soz), die vergleichsweise selten zur Indexierung von Dokumenten in SOLIS verwendet werden. Beispielsweise liefert die Schlagwortsuche nach den Deskriptoren „Biologieunterricht“, „Ethikunterricht“ oder „Pharmakologie“ in SOLIS kaum Treffer, während in der Datenbank FIS Bildung sehr viele und in Psyndex zumindest einige zusätzliche Treffer gefunden werden (siehe Deskriptoren in Fettdruck in Tabelle 6). **Die beiden Deskriptoren „Biologieunterricht“ und „Ethikunterricht“ müssen im Fall der Psyndex Terms durch Kombination von zwei Deskriptoren („Ethik“ + „Lehren“ bzw. „Naturwissenschaftlicher Unterricht“ + „Biologie“) ausgedrückt werden.** Das Beispiel des TheSoz-Deskriptors „schulpsychologischer Dienst“ zeigt, dass die Kombination zweier Deskriptoren nicht immer positive Effekte für eine Konkordanz bringt. Die Kombination der beiden Psyndex-Deskriptoren „Schulpsychologie“ + „Soziale Dienste“ bringt im Fall Psyndex keine Treffer, was beim Scope dieser Datenbank überrascht. Recherchiert man lediglich mit dem Psyndex-Deskriptor „Schulpsychologie“, liefert Psyndex 280 Treffer. Folglich werden durch die Kombination mit dem Deskriptor „Soziale Dienste“ relevante Dokumente ausgeschlossen. Dieser negative Effekt intellektuell erstellter Crosskonkordanzen, insbesondere bei Termkombinationen, kann letztlich nur ausgeschlossen werden, wenn Relationen neben der semantischen Prüfung zusätzlich über den Recall in der Zieldatenbank geprüft werden.

Tabelle 6: Auswahl selten verwendeter TheSoz-Deskriptoren und durch Crosskonkordanzen transformierte Terme in die Vokabulare Thesaurus Bildung und Psyndex Term sowie die Trefferanzahl der Deskriptoren bzw. Deskriptor-Kombinationen in den Datenbanken SOLIS, FIS Bildung & Psyndex (Trefferzahlen sind vom Stand Anfang 2005)

| Thesaurus Sozialwissenschaften | SOLIS | Thesaurus Bildung | FIS Bildung | Psyndex Terms | Psyndex |
|------------------------------------|-------|-------------------------------------|-------------|--|---------|
| Berufsgewerkschaft | 2 | Gewerkschaft | 2175 | Berufe + Gewerkschaften | 2 |
| <i>Biologieunterricht</i> | 7 | <i>Biologieunterricht</i> | 4815 | <i>Naturwissenschaftlicher Unterricht + Biologie</i> | 10 |
| <i>Ethikunterricht</i> | 5 | <i>Ethikunterricht</i> | 818 | <i>Ethik + Lehren</i> | 9 |
| Facharbeiterberuf | 71 | Facharbeiter | 853 | Facharbeiter + Berufe | 2 |
| frauenspezifische Verfolgung | 16 | Frau + Verfolgung | 3 | Verfolgung + Frauen | 2 |
| Islamismus | 64 | Islamischer Fundamentalismus | 4 | Islam + Fundamentalismus | 3 |
| moralisches Urteil | 23 | Moralisches Urteil | 12 | Moral + Beurteilung | 65 |
| nichteheliches Kind | 104 | Uneheliches Kind | 4 | Uneheliche Kinder | 20 |
| <i>Pharmakologie</i> | 8 | <i>Pharmazie</i> | 80 | <i>Pharmakologie</i> | 52 |
| <i>Piaget, J.</i> | 147 | <i>Piaget, Jean</i> | 230 | <i>Piaget (Jean)</i> | 208 |
| Psychometrie | 36 | Psychometrie | 74 | Psychometrie | 438 |
| <i>schulpsychologischer Dienst</i> | 13 | <i>Schulpsychologische Beratung</i> | 306 | <i>Schulpsychologie + Soziale Dienste</i> | 0 |
| Softwareergonomie | 14 | Softwareergonomie | 10 | Software + Ergonomie | 194 |
| Täter-Opfer-Beziehung | 189 | Taeter-Opfer-Beziehung | 14 | Täter + Kriminalitätsopfer | 77 |
| Umweltethik | 49 | Oekologische Ethik | 2 | Umwelt + Ethik | 18 |

Dass auch das Problem der Eigennamen – und insbesondere der mangelnden Standardisierung dieser – durch Crosskonkordanzen positiv beeinflusst werden kann, wird am Beispiel des TheSoz-Deskriptors „Piaget, J.“ deutlich. Durch die zwischengeschaltete Crosskonkordanz können – obwohl alle drei Thesauri den Namen unterschiedlich ansetzen – in allen drei Datenbanken Publikationen des Entwicklungspsychologen „Jean Piaget“ gefunden werden (siehe in diesem Zusammenhang das VIAF-Projekt in (Bennett et al., 2006)).

Die weiteren, in Tabelle 7 aufgezeigten Beispiele sind dem Systematikbereich „3.4 Medizin und Gesundheitswesen“ des Thesaurus Sozialwissenschaften entnommen.

Tabelle 7: TheSoz-Deskriptoren und durch Crosskonkordanzen transformierte Terme des Systematikbereichs 3.4 „Medizin und Gesundheitswesen“ (Trefferzahlen: Stand August 2005)

| Thesaurus Sozialwissenschaften | SOLIS | Thesaurus Bildung | FIS Bildung | Psyndex Terms | Psyndex |
|------------------------------------|-------|--------------------------|-------------|--|---------|
| Krankheit | 6450 | Krankheit | 1506 | Erkrankungen | 2998 |
| öffentliches Gesundheitswesen | 239 | Gesundheitswesen | 628 | Gesundheitsversorgungssystem | 1835 |
| Arzt-Patient-Beziehung | 485 | - | - | Ärzte + Patienten | 234 |
| <i>Ethnomedizin</i> | 72 | - | - | <i>Volksmedizin</i> | 204 |
| körperliche Entwicklung | 75 | Koerperliche Entwicklung | 160 | Physische Entwicklung | 109 |
| gesetzliche Krankenversicherung | 491 | Krankenversicherung | 144 | Krankenversicherung von Arbeitnehmern | 15 |
| <i>Knochenkrankheit</i> | - | <i>Knochenkrankheit</i> | 2 | <i>Osteoporose</i> | 13 |
| Impfung | 7 | Impfung | 33 | Immunisierung | 12 |
| <i>Kinderlähmung</i> | 5 | <i>Kinderlaehmung</i> | 7 | <i>Poliomyelitis</i> | 7 |
| <i>geriatrische Rehabilitation</i> | 1 | - | - | <i>Geriatrische Patienten + Rehabilitation</i> | 123 |

Anhand einiger Deskriptoren (z. B. „Ethnomedizin“, „Knochenkrankheit“, „Kinderlähmung“ und „geriatrische Rehabilitation“), die in den Sozialwissenschaften sicherlich ausgesprochene Spezialgebiete repräsentieren und für die z. T. im Thesaurus Bildung keine Entsprechungen gefunden werden konnten, wird sehr gut deutlich, warum die Crosskonkordanzen in Verbindung mit der interdisziplinären Suche eine entscheidende Verbesserung für den Nutzer bewirken. Für die oben genannten Deskriptoren, die in der Datenbank SOLIS kaum Treffer generieren, bringt gerade die Crosskonkordanz zur Psychologie eine große Erhöhung der Trefferzahlen, die einem sozialwissenschaftlichen Nutzer als Mehrwertleistung angeboten werden kann.

3.2.2.2 Indirekte Term-Transformation

Es ist nicht immer möglich und wünschenswert, sämtliche kontrollierten Vokabulare untereinander zu verknüpfen, da die Erstellung von Crosskonkordanzen – je nach Größe der Vokabulare – mit nicht zu vernachlässigendem Aufwand verbunden ist. Dies gilt nicht nur für fachlich weit entfernte Thesauri in Disziplinen wie z. B. Physik und Psychologie, auch innerhalb eines Faches kann der Aufwand zu groß sein. Beispielsweise besteht in den Sozialwissenschaften der Wunsch, die Datenbank SOLIS der GESIS gemeinsam mit der Datenbank SoLit des Deutschen Zentralinstituts für soziale Fragen (DZI) und dem

Bibliothekskatalog des Sondersammelgebiets Sozialwissenschaften an der Universitäts- und Staatsbibliothek (USB) Köln zu durchsuchen. Letzterer ist mit der Schlagwortnormdatei (SWD) der Deutschen Nationalbibliothek und SoLit mit dem Thesaurus des DZI inhaltlich erschlossen. Der Aufwand, jedes der drei Vokabulare paarweise miteinander zu verbinden, ist bereits relativ groß; zieht man in Betracht, dass dies nur ein kleiner Teil der in den Sozialwissenschaften relevanten Datenbanken ist, so wird schnell deutlich, dass eine paarweise bilaterale Verknüpfung von Vokabularen nicht skaliert, mit vernünftigem Aufwand also nicht zu leisten ist.

Gelöst werden kann dieses Problem durch einen indirekten Term-Transfer, indem bei der Term-Transformation ein „Umweg“ – oder Mehrschrittverfahren – gewählt wird. Dabei wird versucht, durch sequentielle Anwendung mehrerer Crosskonkordanzen einen Transfer zwischen zwei Erschließungssystemen zu erreichen, die nicht direkt miteinander verbunden sind (siehe zur Verdeutlichung das Netz der Crosskonkordanzen in Abbildung 3-5). **Tabelle 8 zeigt eine beispielhafte Anwendung von indirekten Term-Transformationen über eine „switching language“.** Der Thesaurus Sozialwissenschaften (TheSoz) in der mittleren Spalte verbindet die Terme des Startvokabulars (sportwissenschaftliches Vokabular BISp) mit den englischsprachigen Termen des sportwissenschaftlichen Vokabulars (CSA-PAIS). Tabelle 8 beschränkt sich auf die Präsentation der Äquivalenz-Relationen.

Tabelle 8: Indirekte Term-Transformationen. ÄQ-Relationen, BISp->TheSoz (1. TT), TheSoz->CSA-PAIS (2. TT) ausgehend vom Switching Term

| Startterm (BISp) | Switching Term (TheSoz) | Endterm (CSA-PAIS) |
|------------------------|-------------------------|-----------------------|
| Bildung, politische | politische Bildung | Politics + Education |
| Analyse, statistische | statistische Analyse | Statistics + Analysis |
| Presserecht | Medienrecht | Media |
| Primarstufe | Grundschule | Elementary schools |
| Rauschgift | Droge | Drugs |
| Ernährungsstörung | Essstörung | Bulimia |
| Gesundheit, psychische | psychische Gesundheit | Mental health |

Im Anhang „Indirekte Term-Transformationen“ finden sich weitere Termbeispiele für indirekte Term-Transformationen zwischen folgenden kontrollierten Vokabularen:

- Sozialwissenschaften (TheSoz): Switching Language: Gerontologie (Gerolit): Medizin (MeSH)
- Sozialwissenschaften (TheSoz): Switching Language: Psychologie (Psyndex): Medizin (MeSH)
- Sozialwissenschaften (TheSoz): Switching Language: Internationale Beziehungen und Länderkunde (IBLK): Politikwissenschaft (CSA-PAIS)

- Wirtschaftswissenschaften (STW): Switching Language: Internationale Beziehungen und Länderkunde (IBLK): Politikwissenschaft (CSA-PAIS)
- Sozialwissenschaften (CSA-SA): Switching Language: Sozialwissenschaften (TheSoz): Wirtschaftswissenschaften (STW)

Die empirische Fundierung und praktische Erprobung dieses Mehrschrittverfahrens stellt daher auch einen zukünftigen Forschungsschwerpunkt dar.

Nachdem in den vorherigen Abschnitten Beispiele der Erstellung von Crosskonkordanzen vorgestellt wurden, soll im Folgenden auf die Ergebnisse des Projekts KoMoHe eingegangen werden.

3.2.3 Ergebnisse der Mapping-Initiative

Bis Ende 2007 wurden 25 kontrollierte Vokabulare aus 11 Disziplinen und 3 Sprachen (Deutsch, Englisch und Russisch) miteinander verbunden, wobei die Vokabulargrößen zwischen 1.000 - 17.000 abgebildeter Terme pro Vokabular betrugen (Abbildung 3-4 zeigt einen detaillierten Überblick über die Mappings). Mehr als 513.000 Relationen wurden in 65 Crosswalks erstellt (30 bilaterale⁵⁰ und 5 einfache Crosskonkordanzen). Abbildung 3-4 stellt das erstellte Netz an Crosskonkordanzen pro Disziplin dar.

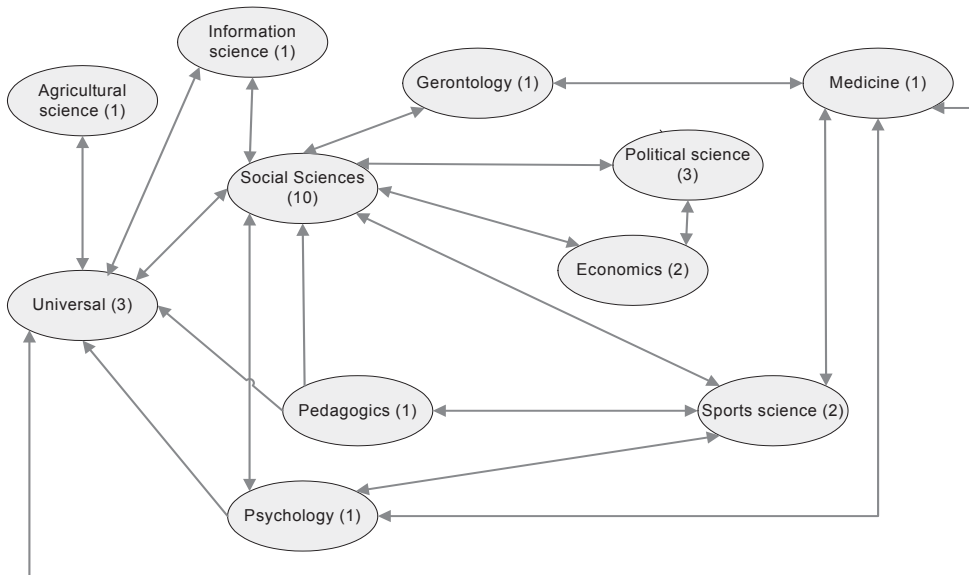


Abbildung 3-4: Netz der Terminologie-Mappings im KoMoHe-Projekt. Die Zahlen in den Klammern enthalten die Anzahl der abgebildeten kontrollierten Vokabulare einer Disziplin.

⁵⁰ Eine bilaterale Crosskonkordanz wird als zwei Crosswalks gezählt.

Das Projekt hat Crosskonkordanzen zwischen insgesamt 25 kontrollierten Vokabularen (Thesauri, Deskriptorenlisten, Klassifikationen und Subject Headings) erstellt, die alle eine Rolle in den fachspezifischen Kollektionen von vascoda spielen. Mehrere Crosskonkordanzen der Vorgängerprojekte CARMEN⁵¹ und infoconnex wurden zudem in das Netz integriert.

Tabelle 9 listet die 25 kontrollierten Vokabulare auf, die über Crosskonkordanzen miteinander verbunden wurden (weitere Details zu den Vokabularen finden sich im Anhang „Kontrollierte Vokabulare“). Im Projekt wurden unterschiedliche Typen von kontrollierten Vokabularen untereinander relationiert: insgesamt 16 Thesauri, 4 Deskriptorlisten, 3 Klassifikationen, 2 Subject Headings. Teilweise wurden aufgrund des Umfangs und wegen fachlicher Gründe nur Ausschnitte aus den Vokabularen verbunden.

Tabelle 9: Kontrollierte Vokabulare im Projekt KoMoHe

| Vokabular-Kürzel | Vokabular |
|------------------|--|
| 1 TheSoz | Thesaurus Sozialwissenschaften |
| 2 Bildung | Thesaurus Bildung |
| 3 Psyndex | Psyndex Terms |
| 4 DZI | Thesaurus of the Deutsches Institut für soziale Fragen |
| 5 CSA-SA | Thesaurus of Sociological Indexing Terms |
| 6 MeSH | Medical Subject Headings |
| 7 BISp | Descriptors of the Bundesinstitut für Sportwissenschaft |
| 8 STW | Standard Thesaurus Wirtschaft |
| 9 IBLK | Thesaurus Internationale Beziehungen und Länderkunde (Euro-Thesaurus) |
| 10 FES | Descriptors of the Friedrich-Ebert-Stiftung |
| 11 GEROLIT | Thesaurus of the Deutsches Zentrum für Altersfragen |
| 12 SWD | Schlagwortnormdatei |
| 13 CSA-WPSA | CSA Thesaurus of Political Science Indexing Terms |
| 14 CSA-PEI | CSA Thesaurus Physical Education Index |
| 15 CSA-ASSIA | CSA Thesaurus Applied Social Sciences Index and Abstracts |
| 16 TWSE | Thesaurus für wirtschaftliche und soziale Entwicklung |
| 17 ELSST | European Language Social Science Thesaurus |
| 18 INFODATA | INFODATA Thesaurus |
| 19 IAB | Descriptors of the Institut für Arbeitsmarkt- und Berufsforschung |
| 20 AGROVOC | AGROVOC Thesaurus |
| 21 JEL | Journal of Economic Literature Classification System |
| 22 RVK | Regensburger Verbundklassifikation |
| 23 DDC | Dewey Decimal Classification |
| 24 CSA-PAIS | CSA Thesaurus PAIS International Subject Headings |
| 25 INION | Descriptors of the Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences |

51 <http://www.bibliothek.uni-regensburg.de/projects/carmen12/index.html.en>

Ein Großteil der Vokabulare stammt aus den Sozialwissenschaften und verwandten Disziplinen wie Politikwissenschaft, Pädagogik usw. Es wurden zudem zahlreiche Vokabulare aus nicht-sozialwissenschaftlichen Disziplinen wie z.B. Agrarwissenschaft, Sport, Informationswissenschaft, Wirtschaftswissenschaften, Psychologie und Medizin bearbeitet. Des Weiteren wurden fachspezifische Ausschnitte der Universalvokabulare (z.B. SWD, RVK, DDC) bearbeitet (siehe dazu Abbildung 3-4).

Die im Projekt KoMoHe verwendeten Vokabulare sind hauptsächlich in Deutsch, Englisch (N=8), Russisch (N=1) oder multilingual (z.B. AGROVOC, IBLK, DDC). Einige Vokabulare beinhalten englische oder deutsche Übersetzungen der Terme (z.B. THESOZ, PSYINDEX, MESH, INION, STW). Die folgende Auflistung unterteilt die kontrollierten Vokabulare im Projekt KoMoHe in Thesauri, Deskriptorenlisten, Klassifikationen und Subject Headings und beschreibt die Vokabulare.

Involvierte Thesauri (N=16):

- AGROVOC Thesaurus (AGROVOC): A vocabulary in the *agricultural* domain which contains round 39,000 terms. Mapping to: SWD.
- CSA Thesaurus Applied Social Sciences Index and Abstracts (CSA-ASSIA): A vocabulary in the *social science* domain which contains round 17,000 terms. Mapping to: THESOZ.
- CSA Thesaurus PAIS International Subject Headings (CSA-PAIS): A vocabulary in the *political science* domain which contains round 7,000 terms. Mapping to: IBLK.
- CSA Thesaurus Physical Education Index (CSA-PEI): A vocabulary in the *sports science* domain which contains round 1,800 terms. Mapping to: THESOZ.
- CSA Thesaurus of Political Science Indexing Terms (CSA-WPSA): A vocabulary in the *social and political science* domain which contains round 3,100 terms. Mapping to: THESOZ.
- European Language Social Science Thesaurus (ELSST): A vocabulary in the *social science* domain which contains round 3,200 terms. Mapping to: THESOZ.
- INFODATA Thesaurus (INFODATA): A vocabulary in the *information science* domain which contains round 1,000 terms. Mapping to: THESOZ and SWD.
- Psynindex Terms (PSYINDEX): A vocabulary in the *psychological* domain which contains round 5,400 terms. Mapping to: THESOZ, SWD, BISP, MESH and BILDUNG.
- Standard Thesaurus Wirtschaft (STW): A vocabulary in the *economics* domain which contains round 5,700 terms. Mapping to: THESOZ, SWD, IAB and IBLK.
- Thesaurus Bildung (BILDUNG): A vocabulary in the *pedagogic* domain which contains round 50,000 terms. Mapping to: THESOZ, SWD, PSYINDEX and BISP.
- Thesaurus Internationale Beziehungen und Länderkunde (IBLK): A vocabulary in the *political science* domain which contains round 8,400 terms. Mapping to: THESOZ, STW, TWSE and CSA-PAIS.
- Thesaurus Sozialwissenschaften (THESOZ): A vocabulary in the *social science* domain which contains round 7,700 terms. Mapping to: GEROLIT, DZI, FES, CSA-WPSA, CSA-ASSIA, CSA-SA, CSA-PEI, ELSST, IAB, IBLK, STW, SWD, BILDUNG, PSYINDEX, INFODATA and BISP.

- Thesaurus für wirtschaftliche und soziale Entwicklung (TWSE): A vocabulary in the *political science* domain which contains round 2,800 terms. Mapping to: IBLK.
- Thesaurus of Sociological Indexing Terms (CSA-SA): A vocabulary in the *social science* domain which contains round 4,300 terms. Mapping to THESOZ.
- Thesaurus of the Deutsches Institut für soziale Fragen (DZI): A vocabulary in the *social science* domain which contains round 1,900 terms. Mapping to THESOZ.
- Thesaurus of the Deutsches Zentrum für Altersfragen (GEROLIT): A vocabulary in the *gerontology* domain which contains round 1,900 terms. Mapping to THESOZ and MESH.

Involvierte Deskriptorenlisten (N=4):

- Descriptors of the Bundesinstitut für Sportwissenschaft (BISP): A vocabulary in the *sports science* domain which contains round 7,400 terms. Mapping to THESOZ, MESH and BILDUNG.
- Descriptors of the Friedrich-Ebert-Stiftung (FES): A vocabulary in the *social science* domain which contains round 4,000 terms. Mapping to THESOZ.
- Descriptors of the Institut für Arbeitsmarkt- und Berufsforschung (IAB): A vocabulary in the *social science* domain which contains round 6,800 terms. Mapping to THESOZ and STW.
- Descriptors of the Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences (INION): A vocabulary in the *social science* domain which contains round 7,000 terms. Mapping to THESOZ.

Involvierte Klassifikationen (N=3):

- Dewey Decimal Classification (DDC): A *universal* vocabulary which contains thousands of notations. Mapping to RVK.
- Journal of Economic Literature Classification System (JEL): A vocabulary in the *economics* domain which contains round 1,000 notations. Mapping to STW.
- Regensburger Verbundklassifikation (RVK): A *universal* vocabulary which contains thousands of notations. Mapping to DDC.

Involvierte Subject Headings (N=2):

- Medical Subject Headings (MESH): A vocabulary in the *medicine* domain which contains round 23,000 terms. Mapping to PSYINDEX, GEROLIT, BISP and SWD.
- Schlagwortnormdatei (SWD): A *universal* vocabulary which contains round 650,000 terms. Mapping to THESOZ, MESH, STW, AGROVOC and INFODATA.

Abbildung 3-5 gibt einen Überblick über alle 65 Crosskonkordanzen. Der Thesaurus Sozialwissenschaften (THESOZ) ist das Vokabular mit den meisten ein- und ausgehenden Verbindungen, und aufgrund seiner Zentralität wird der THESOZ in der Mitte des Netzes dargestellt. Andere Vokabulare wie die SWD oder der PSYINDEX spielen eine zentrale Rolle, um in andere Fachdomänen zu wechseln (vgl. (Bates, 1998)). Die Crosskonkordanz DDC-RVK ist die einzige Crosskonkordanz, die nicht mit dem Netz verbunden ist. Möglicherweise kann die Terminologearbeit, die im Projekt CRISSCROSS zwischen SWD

und DDC entstanden ist, genutzt werden, um dieses unverbundene Paar in das Netz einzugliedern. Die Konkordanz JEL-STW ist ein Beispiel für eine einfache (one-way) Crosskonkordanz, ausgehend von JEL zu STW.

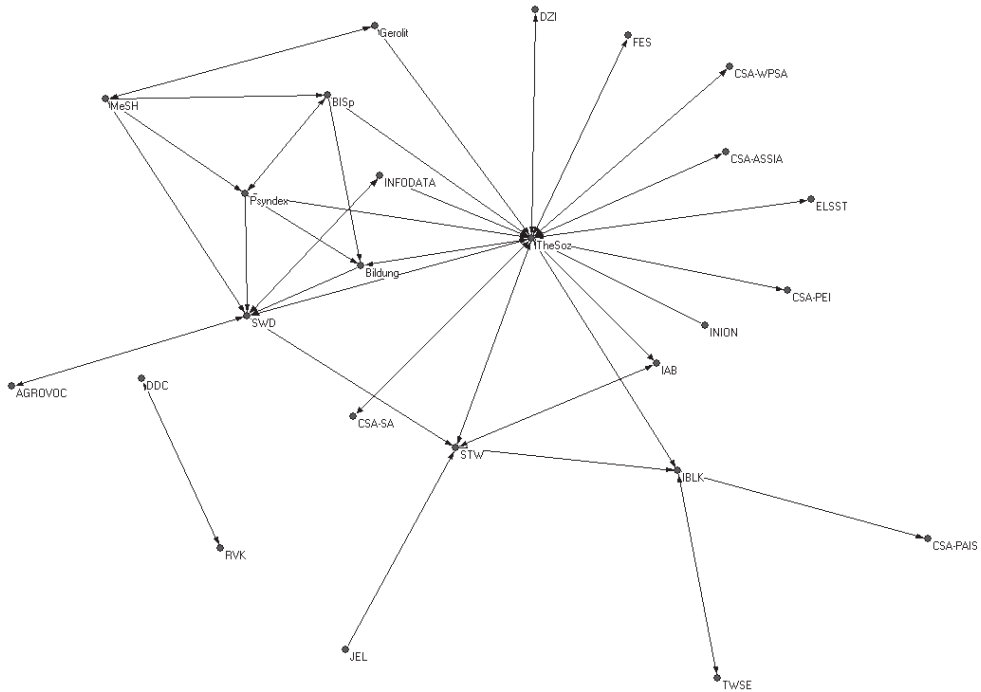


Abbildung 3-5: Netz der über Crosskonkordanzen verbundenen Vokabulare im KoMoHe-Projekt

Die 513.000 Relationen, die in der Crosskonkordanz-Datenbank der GESIS verfügbar sind, beinhalten mehr als 181.000 einzelne suchbare Konzepte (einzelne kontrollierte Deskriptoren oder Deskriptor-Kombinationen oder Notationen). Im Durchschnitt (pro Crosskonkordanz) werden 6.500 Startterme zu 3.600 Termen im Endvokabular verbunden (1,2 Relationen pro Term).

Tabelle 10 zeigt die Crosskonkordanzen, die im Projekt KoMoHe von 2004 bis 2007 erarbeitet wurden (Nr. 1-27). Zudem enthält Tabelle 10 **fünf bilaterale und drei unilaterale** Crosskonkordanzen, die in den Vorgängerprojekten vor 2004 erstellt wurden. Bis auf wenige Ausnahmen wurden alle Crosskonkordanzen bilateral erstellt.

Tabelle 10: Übersicht der Crosskonkordanzen im Projekt KoMoHe und Vorgängerprojekten

| | Disziplin A | Disziplin B | Vok. A | Vok. B | Typ | Jahr |
|----|-------------|-------------|-----------------|------------------|-----------|----------|
| 1 | Soz | Wirtschaft | <i>TheSoz</i> | <i>STW</i> | bilateral | 2004 |
| 2 | Soz | Sport | <i>TheSoz</i> | <i>BiSp</i> | bilateral | 2004 |
| 3 | Psy | Sport | <i>Psynder</i> | <i>BISp</i> | bilateral | 2004 |
| 4 | Sport | Päda | <i>BISp</i> | <i>Bildung</i> | | 2004 |
| 5 | Soz | Soz | <i>TheSoz</i> | <i>DZI</i> | bilateral | 2005 |
| 6 | Soz | Pol | <i>TheSoz</i> | <i>FES</i> | bilateral | 2005 |
| 7 | Soz | Pol | <i>TheSoz</i> | <i>IBLK</i> | bilateral | 2005 |
| 8 | Soz | Geron | <i>TheSoz</i> | <i>Gerolit</i> | bilateral | 2005 |
| 9 | Med | Sport | <i>MeSH</i> | <i>BISp</i> | bilateral | 2005 |
| 10 | Wirtschaft | Pol | <i>STW</i> | <i>IBLK</i> | bilateral | 2005 |
| 11 | Soz | Soz | <i>TheSoz</i> | <i>CSA-WPSA</i> | bilateral | 2006 |
| 12 | Soz | Soz | <i>TheSoz</i> | <i>CSA-ASSIA</i> | bilateral | 2006 |
| 13 | Soz | Soz | <i>TheSoz</i> | <i>ELSST</i> | bilateral | 2006 |
| 14 | Soz | Sport | <i>TheSoz</i> | <i>CSA-PEI</i> | bilateral | 2006 |
| 15 | Med | Psy | <i>MeSH</i> | <i>Psynder</i> | bilateral | 2006 |
| 16 | Med | Geron | <i>MeSH</i> | <i>Gerolit</i> | bilateral | 2006 |
| 17 | Pol | Pol | <i>IBLK</i> | <i>CSA-PAIS</i> | bilateral | 2006 |
| 18 | Pol | Pol | <i>IBLK</i> | <i>TWSE</i> | bilateral | 2006 |
| 19 | Infwiss | Allgemein | <i>INFODATA</i> | <i>SWD</i> | bilateral | 2007 |
| 20 | Infwiss | Soz | <i>INFODATA</i> | <i>TheSoz</i> | bilateral | 2007 |
| 21 | Soz | Soz | <i>IAB</i> | <i>TheSoz</i> | bilateral | 2007 |
| 22 | Soz | Econ | <i>IAB</i> | <i>STW</i> | bilateral | 2007 |
| 23 | Allgemein | Med | <i>SWD</i> | <i>MeSH</i> | bilateral | 2007 |
| 24 | Allgemein | Agrar | <i>SWD</i> | <i>AGROVOC</i> | bilateral | 2007 |
| 25 | Soz | Soz | <i>INION</i> | <i>TheSoz</i> | | 2007 |
| 26 | Wirtschaft | Wirtschaft | <i>JEL</i> | <i>STW</i> | | 2007 |
| 27 | Allgemein | Allgemein | <i>RVK</i> | <i>DDC</i> | bilateral | 2007 |
| 28 | Soz | Allgemein | <i>TheSoz</i> | <i>SWD</i> | bilateral | vor 2004 |
| 29 | Soz | Soz | <i>TheSoz</i> | <i>CSA-SA</i> | bilateral | vor 2004 |
| 30 | Soz | Psy | <i>TheSoz</i> | <i>Psynder</i> | bilateral | vor 2004 |
| 31 | Soz | Päda | <i>TheSoz</i> | <i>Bildung</i> | bilateral | vor 2004 |
| 32 | Psy | Päda | <i>Psynder</i> | <i>Bildung</i> | bilateral | vor 2004 |
| 33 | Psy | Allgemein | <i>Psynder</i> | <i>SWD</i> | | vor 2004 |
| 34 | Päda | Allgemein | <i>Bildung</i> | <i>SWD</i> | | vor 2004 |
| 35 | Wirtschaft | Allgemein | <i>STW</i> | <i>SWD</i> | | vor 2004 |

Die Erstellung der Crosskonkordanzen im Projekt konzentrierte sich auf zwei Bereiche (siehe (Mayr und Walter, 2008)):

- Crosskonkordanzen innerhalb eines Faches
- Crosskonkordanzen zwischen Fächern

Es wurde versucht, ein Fach anhand der verfügbaren Vokabulare möglichst vollständig zu vernetzen. Dabei wurden sowohl innerdisziplinäre als auch interdisziplinäre Verbindungen angestrebt.

Am Beispiel des Faches Sport, aber auch der Sozialwissenschaften lässt sich die unterschiedliche Vorgehensweise illustrieren. Während die sportwissenschaftliche Deskriptorenliste des Bundesinstituts für Sportwissenschaft (BISp) zu den Fächern Psychologie (Psyndex), Medizin (MeSH), Pädagogik (Bildung) und den Sozialwissenschaften (TheSoz), also jeweils zu fremden Disziplinen (interdisziplinär) verknüpft wurde, wurde der Thesaurus Sozialwissenschaften häufiger innerdisziplinär verbunden (sozialwissenschaftliche Vokabulare: DZI, CSA-SA, FES, CSA-ASSIA, CSA-WPSA, ELSST).

Das Fach Sport bzw. das Sportportal sport-if⁵² hat dabei die Möglichkeit, auf Basis der erstellten Crosskonkordanzen Suchanfragen an das Medizinportal medpilot (MeSH-Crosswalk), das Psychologieportal⁵³ (Psyndex-Crosswalk), das Fachportal Pädagogik⁵⁴ (Bildung-Crosswalk) sowie das sozialwissenschaftliche Portal sowiport (TheSoz-Crosswalk) zu stellen.

Abbildung 3-6 zeigt die Verteilung der Relationstypen im Projekt (vgl. Tabelle 5). Äquivalenzrelationen (ca. 45 %) sind der häufigste Relationstyp zwischen Termen. Lediglich 12 % von allen Relationen sind Null-Relationen (ein Term kann nicht verbunden werden).

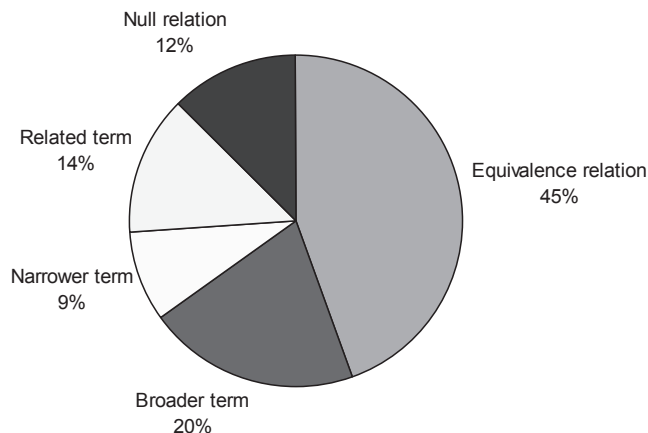


Abbildung 3-6: Verteilung der Relationstypen über alle Crosskonkordanzen.

52 <http://www.sport-if.de/>

53 <http://www.zpid-psychologie.de/>

54 <http://www.fachportal-paedagogik.de/>

Die intrinsischen Merkmale von Crosskonkordanzen (und ihre Auswirkung auf die Suche) können sich abhängig von den abgebildeten kontrollierten Vokabularen und externen Faktoren im Erstellungsprozess⁵⁵ der Crosskonkordanzen unterscheiden. Änderungen in den kontrollierten Vokabularen oder in der Indexierungspraxis können die Qualität der Crosskonkordanz zudem beeinflussen. Folgende Unterschiede bei der Erstellung von Crosskonkordanzen konnten im Projekt KoMoHe beobachtet werden:

- Größe der Start- und Zielvokabulare
- Unterschiede im Grad der Prä- und Postkoordination der Vokabulare
- Anzahl der Relationen
- Anzahl der abgebildeten Zielterme (Abdeckung/Überlappung)
- Verteilung der Relationen (Äquivalenz, breiterer Term, engerer Term, verwandter Term, Null-Relation)
- Verteilung der Relevanzen (hoch, mittel, gering)
- identische Terme
- Unterschiede in der Spezifität (z. B. Vokabulare mit einem sehr allgemeinen oder spezifischen Scope)
- Kombinationen von abgebildeten Termen (Mappings, die aus mehr als einem Endterm bestehen).

Eine quantitative Analyse kann einige Aufschlüsse über die Grundmerkmale einer Crosskonkordanz ergeben, die Qualitätsverbesserungen durch die spezifischen Mappings für die Suche können damit aber nicht bestimmt werden (siehe dazu Abschnitt 5.2.1).

In den nächsten Abschnitten wird die Implementation der Crosskonkordanzen beschrieben.

3.2.4 Implementation der Crosskonkordanzen

Um die Crosskonkordanzen für den späteren Gebrauch zu speichern, wurde eine relationale Datenbank erstellt. Die Datenbank wurde von Anne-Kathrin Walter konzipiert und aufgesetzt. Es stellte sich heraus, dass die relationale Struktur geeignet war, die Menge der unterschiedlichen kontrollierten Vokabulare, Terme, Kombinationen von Termen und Beziehungen angemessen abzubilden. Die Vokabulare und Terme werden in Listenform repräsentiert, wobei diese unabhängig voneinander und ohne Berücksichtigung der syntetischen Struktur der involvierten Vokabulare gespeichert werden. Orthographie und Groß-/Kleinschreibung der Terme der kontrollierten Vokabulare werden normalisiert.

⁵⁵ Zum Beispiel kann das Erstellungsdatum der Crosskonkordanz die Anzahl der Relationen pro Startterm beeinflussen. Im Projekt wurden beispielsweise zu Beginn weniger Relationen erstellt als zum Ende des Projekts. Die Crosskonkordanzen des Vorgängerprojekts (CARMEN) wurden z. T. in der Expertengruppe intensiv diskutiert und deutlich selektiver erstellt als im KoMoHe-Projekt (siehe dazu CARMEN-Projekt (2002): CARMEN - Abschlussbericht des Arbeitspakets 12 (AP 12) Crosskonkordanzen von Klassifikationen und Thesauri. URL: http://www.opus-bayern.de/uni-regensburg/volltexte/2003/242/pdf/CARMENAP12_Abschlussbericht_Netz.pdf).

Kombinationen aus Termen (z. B. „computers“ + „crime“ als verwandte Kombination für den Term „hacker“ in Tabelle 5) werden als separate Konzepte gespeichert.

Um die Terminologie-Daten in der Datenbank zugänglich zu machen, wurde ein Web Service, der sog. Heterogenitätsservice oder HTS entwickelt (siehe dazu (Hill et al., 2002, Mayr und Walter, 2008)), der die Crosskonkordanz-Suche nach individuellen Starttermen, abgebildeten Termen, Start- und Zielvokabularen sowie den unterschiedlichen Relationstypen unterstützt. Eine Implementation des Heterogenitätsservice, die ausschließlich Äquivalenzrelationen nutzt, schlägt Suchterme in den Termlisten der kontrollierten Vokabulare nach und fügt alle äquivalenten Suchterme aus allen zugänglichen Vokabularen zu der ursprünglichen Suche. Wenn die kontrollierten Vokabulare in unterschiedlichen Sprachen vorliegen, bietet der Heterogenitätsservice eine Übersetzung des Originalterms zu dem bevorzugten kontrollierten Term in der anderen Sprache. Auch wenn die ursprüngliche Anfrage einen booleschen Operator enthält, bleibt dieser nach der Anfrageerweiterung intakt (jeder Suchbegriff wird separat transformiert).

Aufgrund von Performance-Gründen wird die Suchterm-Erweiterung durch Crosskonkordanzen im Suchsystem sowiport nicht für jede einzelne Datenbank und ihr entsprechendes kontrolliertes Vokabular unterschieden, es werden aber alle äquivalenten Terme zu der Anfrage hinzugefügt. Im Prinzip erweitert die Benutzung des Terminologie-Netzes eine beliebige Term-Anfrage mit Synonymen oder Quasi-Synonymen (siehe dazu unteres Beispiel in Abbildung 3-7). In (Mayr und Petras, 2008) finden sich weitere Überlegungen zur Implementation der Crosskonkordanzen zur Unterstützung der Suche (siehe auch (Stempfhuber, 2003)).

“By adding terms to the query, recall should increase, that is, more relevant documents will be found. It is unclear, however, whether the indiscriminate expansion of the original query without regard for the terms’ appropriateness for a given database can actually decrease the precision of the search. If the created equivalence mappings only denote correct synonyms, then the adding of true synonyms should have no such effect. However, homonymic terms as well as slight variations in the meaning of a concept can have a detrimental impact on the quality and precision of the query. In an ideal case, the searcher could be represented with a selection of terms garnered from the cross-concordances and then select an appropriate formulation.” (Mayr und Petras, 2008)



Abbildung 3-7: Benutzeranzeige von Termtransformation in sowiport. Die transformierten Terme werden bei der Abfrage über ein Boolesches OR verknüpft und an die Datenbanken geschickt.

3.2.5 Heterogenitätsservice (HTS)

Die erstellten Crosskonkordanzen werden über den Heterogenitätsservice (HTS) verfügbar gemacht. Die erste Version des HTS wurde von Anne-Kathrin Walter im Projekt konzipiert und implementiert. Die aktuelle Version des HTS, der über sowiport und künftig vascoda angesprochen wird, wurde von Stefan Bärish entwickelt. Maximilian Stempfhuber leitete die technische Arbeitsgruppe im Projekt KoMoHe. In diesem Abschnitt wird anhand eines Einsatzszenarios die Funktionalität der ersten Version des HTS vorgestellt und die Datenbasis beschrieben, auf die er zugreift und die gleichzeitig das Speicherformat der Crosskonkordanzen ist.

Es gibt mehrere Einsatzmöglichkeiten für den HTS. Basisfunktionalität ist der Dienst des Term-Transfers für Fachportale als reine Maschinen-Schnittstelle (vgl. (Binding und Tudhope, 2004)). Weiterhin ist ein Einsatz des Service als Rechercheunterstützung für den Nutzer denkbar. Das durch die Crosskonkordanzen entstandene Terminologie-Netz kann bei der Formulierung seines Suchbedürfnisses hilfreich sein. Ferner könnte der Service in Zukunft Funktionen zum Update der Konkordanzen umfassen. Anhand des im Folgenden beschriebenen Szenarios werden Entscheidungen zur technischen Realisierung, zur Schnittstelle und zur Architektur des Service erläutert.

Ein Nutzer hat ein Informationsbedürfnis und formuliert seine Anfrage in dem ihm vertrauten kontrollierten Vokabular A (Ausgangsvokabular), das Grundlage der inhaltlichen Erschließung in Datenbank A ist. Die Datenbanken B und C sind mit anderen kontrollierten Vokabularen erschlossen. Ziel des Fachportals, das die drei Datenbanken zur integrierten Recherche anbietet, ist es, dem Nutzer möglichst viele relevante Dokumente bezogen auf sein Informationsbedürfnis zu liefern. Bevor es die Anfrage an die Datenbanken weitergibt, wird der HTS nach Transformationen in die Vokabulare (Zielvokabulare) der Datenbanken B und C gefragt. Falls vorhanden, wird die Anfrage pro Datenbank modifiziert und anschließend die Abfrage gestartet.

Abbildung 3-8 verdeutlicht das Szenario.

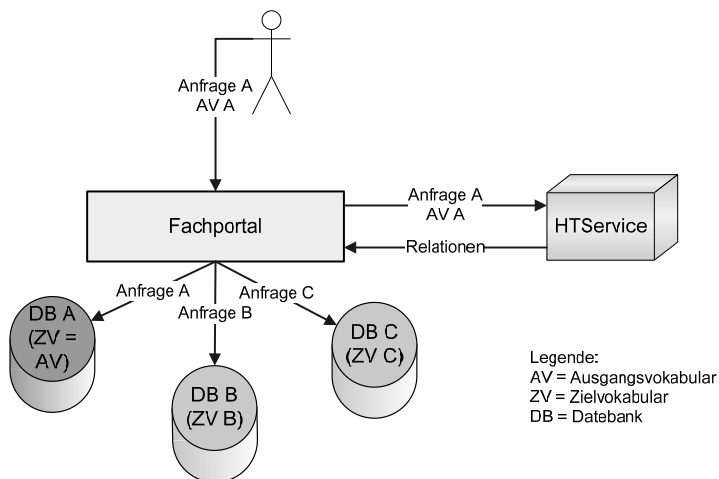


Abbildung 3-8: Einsatzszenario des Heterogenitätsservice (HTS)

Grundlage für den HTS ist die Web Service-Technologie. Das Kommunikationsprotokoll SOAP⁵⁶ als deren Basis ermöglicht es, dass Fachportal und HTS unabhängig vom unterliegenden Übertragungsprotokoll und lokal verwendeten Technologien kommunizieren. Da SOAP ein XML-basiertes Protokoll ist, bleibt die Kommunikation für Menschen lesbar, ist aber auch für Maschinen prozessierbar. Zudem ist es ein offener Standard, der ohne Einschränkungen zugänglich ist. Die Realisierung als Web Service bietet einen weiteren Vorteil für die automatisierte Kommunikation zwischen Anwendungen: Es gibt ein standardisiertes Format zur Beschreibung der Schnittstelle, d. h. es ist spezifiziert, welche Funktionen der Service anbietet, wie sie aufgerufen werden und wie die Antwort aufgebaut ist. Auf diese Weise kann sehr einfach eine Anfrage an den Dienst erfolgen. Die Anfrage eines Fachportals an den HTS kann je nach Suchanfrage des Nutzers unterschiedlich strukturiert sein. Immer enthalten ist natürlich der Ausgangsterm, der transformiert werden soll. Abhängig von der Suche, die der Nutzer durchführt, können weitere Einschränkungen angegeben sein.

- Relationstyp: Wie in Abschnitt 3.2.2 beschrieben, gibt es unterschiedliche Relationstypen, die die Deskriptoren verbinden. Ober- und Unterbegriffsrelationen liefern weitere oder engere transformierte Terme, daher ist davon auszugehen, dass die Treffermenge bezüglich des Ausgangsterms und damit bezüglich des Informationsbedürfnisses des Nutzers zu groß bzw. zu speziell ist. Das gleiche gilt für die Ähnlichkeitsrelation: Sie liefert ein verwandtes Konzept zur ursprünglichen Anfrage. Die beste Abbildung wird durch die Äquivalenzrelation erbracht. Es ist daher empfehlenswert, nur letztere automatisiert einzusetzen und dem Nutzer die weiteren Relationen zur Verfeinerung bzw. Ausweitung seiner Suche anzubieten. Es muss daher möglich sein, die Anfrage an den Heterogenitätsservice auf einen bestimmten Relationstyp einzuschränken.
- Bei der erweiterten Suche kann ein Nutzer die Datenbanken auswählen, in denen er suchen möchte. Durch die Auswahl sind die Zielvokabulare bekannt, in die transformiert werden soll, d. h. die Relationen können bei der Anfrage an den Heterogenitätsservice auf diese eingeschränkt werden.
- Eventuell hat ein Nutzer seine Suchterme aus einem Online-Thesaurus oder Search Term Recommender (Petras, 2006) ausgewählt und auf diese Weise das Ausgangsvokabular, von dem aus transformiert werden soll, vorgegeben. Da Terme in mehreren Vokabularen vorkommen können, sollte auch das Ausgangsvokabular in der Anfrage festgelegt werden können.
- Längerfristig soll der Heterogenitätsservice auch andere Transformationen als die intellektuell erstellten zurückgeben (z. B. durch statistische Verfahren ermittelte Relationen), daher wird in der Anfrage noch ein Feld vorgesehen, in dem die Transformationsmethode spezifiziert werden kann.

Für das Format von Anfrage und Rückgabe wird ebenfalls XML gewählt. Es gelten die gleichen Vorteile: Die Kommunikation ist sowohl durch Anwendungen prozessierbar, aber auch menschenlesbar und XML ist ebenfalls ein offener, frei zugänglicher Standard.

56 <http://www.w3.org/TR/soap12-part1/>

Abbildung 3-9 zeigt das Format der Anfrage. Der Übersichtlichkeit halber ist die Anfrage nicht in XML, sondern als Baumstruktur dargestellt. Die Klammern bedeuten, dass dieser Parameter optional ist.

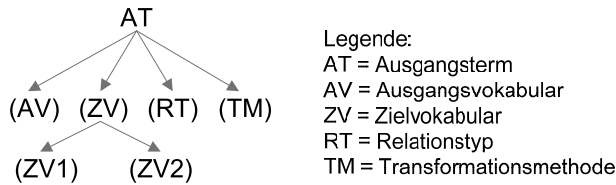


Abbildung 3-9: Format der Anfrage

Um das Auswerten des Ergebnisses zu erleichtern, sollte das Format der Rückgabe einheitlich sein, unabhängig davon, wie viele Einschränkungen (z. B. Zielvokabular, Relationstyp) in der Anfrage spezifiziert wurden. Es ist allerdings nicht ausreichend, nur die transformierten Terme zurückzugeben, da sonst unklar ist, für welches Zielvokabular sie sind. Weiterhin sollte eine Zuordnung von Ausgangs- zu Zielvokabular erfolgen, damit ersichtlich ist, welche Konkordanz angewendet wurde. Für die Rückgabe ergibt sich damit eine Baumstruktur, die anhand des Anfrageterms „Bildungseinrichtung“ in Abbildung 3-10 beispielhaft dargestellt ist.

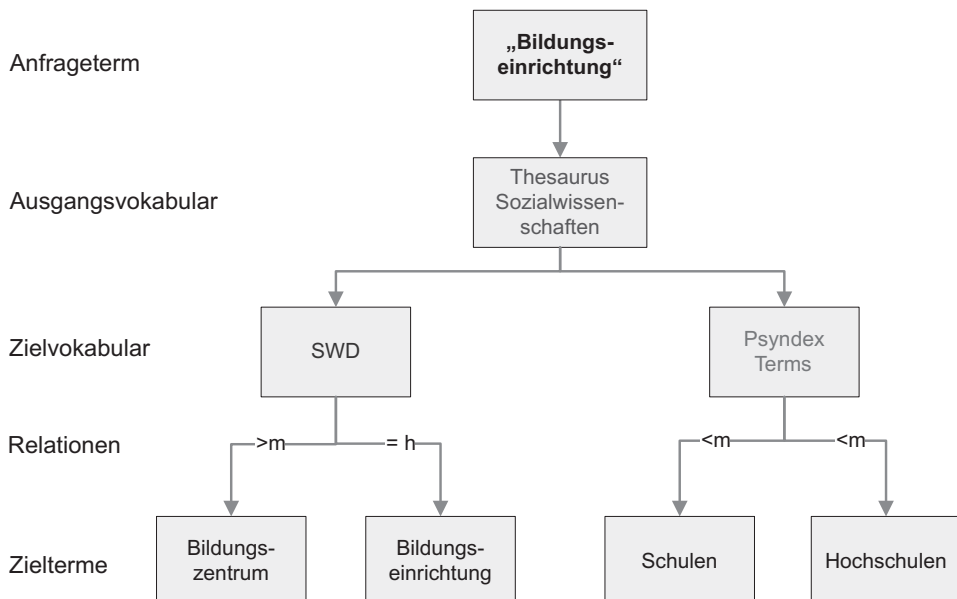


Abbildung 3-10: Beispiel für die Rückgabe

Die Erstellung von Crosskonkordanzen erfolgt in Tabellen, die allerdings als Datenbasis für den Heterogenitätsservice nicht geeignet sind, da sie leicht verändert oder einfach verschoben bzw. gelöscht werden können. Für eine persistente Speicherung, die gleichzeitig einen zuverlässigen Zugriff ermöglicht, bietet es sich an, die Crosskonkordanzen in einer Datenbank abzulegen. Ein weiterer Vorteil davon ist, dass eine Selektierbarkeit und Auswahl der Relationen nach unterschiedlichen Kriterien möglich ist.

Die Speicherung in einer Datenbank erfordert ein Tabellen-Schema, das folgenden Anforderungen genügen muss:

- Kein Informationsverlust gegenüber den Tabellen, in denen die Konkordanzen erstellt werden: Sämtliche Angaben über Relationen, Relevanzen und Zielterme müssen in der Datenbank wiederzufinden sein.
- Selektierbarkeit: Die Crosskonkordanzen sollten nach verschiedenen Kriterien selektierbar sein.

Ausgangsterm: Die Transformation einer Anfrage muss bearbeitet werden können, ohne jede Crosskonkordanz einzeln durchsuchen zu müssen; daher werden alle Relationen in einer einzigen Tabelle abgespeichert. Terme, die aus unterschiedlichen Thesauri kommen, sich aber nur in der Groß-/Kleinschreibung oder hinsichtlich der Schreibweise von Umlauten unterscheiden, müssen ebenfalls durch eine einzelne Abfrage zu ermitteln sein. Neben der Originalschreibweise werden sie daher auch in einer normierten Schreibweise (Großschreibung und ohne Umlaute) vorgehalten.

Ausgangs- und Zielvokabular: Die Speicherung aller Relationen in einer Tabelle erfordert, dass eine Zuordnung von Relation zu Konkordanz möglich ist. Daher wird für jede Transformation Ausgangs- und Zielvokabular in Extraspalten gespeichert. Relationstyp: Da die verschiedenen Relationstypen unterschiedliche Auswirkungen auf die Treffermenge haben, sollte es möglich sein, die Relationen auf einen Typ, z. B. die Äquivalenzrelation, zu begrenzen.

Vor dem Laden in die Datenbank werden sowohl Terme als auch Relationen und Relevanzen auf syntaktische Korrektheit überprüft, d. h. die richtige Schreibweise für die Terme sowie nur die erlaubten Relationen und Relevanzen. Erwähnenswert ist, dass nicht alle Terme eines Thesaurus auch in den Term-Transformations-Tabellen zu finden sind, da zum Teil nur Ausschnitte von Thesauri verknüpft wurden (z. B. sozialwissenschaftlicher Ausschnitt der SWD in der CK TheSoz-SWD).

Weiterhin ist geplant, die Mappings in das SKOS-Format zu konvertieren, um damit den Anschluss an die Semantic Web Community zu gewährleisten (Assem et al., 2006, Hage, 2008, Miles, 2006).

“Aside from its application in a distributed search scenario, the semantic web community might be able to find new and interesting usages for terminology data like this one. The SKOS standard (Simple Knowledge Organization System) contains a section on mapping vocabularies in its draft version. Once the standard gets stabilized, we plan on transferring the cross-concordance data to the SKOS format. If more vocabularies and mappings become available in SKOS, then further research into connecting previously unmapped terminology networks with each other should be possible.” (Mayr und Petras, 2008)

Nachdem in diesem Kapitel die Bemühungen des Projekts KoMoHe beim Zusammenführen von heterogenen Informationsangeboten beschrieben wurden, soll im folgenden Kapitel konkreter auf Bradford's Law of Scattering eingegangen werden, das die Grundlage für das Verfahren Bradfordizing darstellt. Die Evaluation von Bradfordizing wird in Kapitel 5 beschrieben.

4 Bradford's Law of Scattering

Das folgende Kapitel trägt die Grundlagen von Bradford's Law of Scattering zusammen und stellt anschließend die Anwendung von Bradfordizing im Zusammenhang mit dieser Arbeit vor.

Der Begriff Bradford's Law of Scattering (BLS) stammt ursprünglich von B. Vickery (Vickery, 1948), der sich sehr früh mit den Grundlagen der Bradfordschen Gesetzmäßigkeit beschäftigt hat. Die Beobachtung des Bibliothekars Samuel Bradford aus den frühen 1930er Jahren kann als eine der Wurzeln der Informetrie (engl. Informetrics) bezeichnet werden. Die Informetrie ist wiederum eine Spezialdisziplin der Informationswissenschaft, die sich mit der mathematischen Beschreibung und Modellierung der Regularitäten (vgl. (Bookstein, 1990, Yablonsky, 1980)) bzw. sogenannten „Gesetzmäßigkeiten“ der beobachtbaren Objekte im Bibliotheks- und Informationswesen beschäftigt. Zu diesen Objekten zählen z. B. Autoren, Publikationen (Zeitschriften, Monographien usw.), Referenzen, Zitationszahlen usw. Wichtig bei der Beurteilung der informetrischen Gesetzmäßigkeiten ist, dass diese keine Gesetzmäßigkeiten in Sinne von Naturgesetzen wie z. B. der Fallgesetze oder des Gravitationsgesetzes darstellen, sondern empirische Gesetze (Wilkinson, 1972). Diese können zwar mathematisch über Formeln beschrieben werden, Abweichungen und Unschärfe bei der Formulierung und Beobachtung der informetrischen Zusammenhänge sind aber die Norm. Häufig können die informetrischen Modelle die Realität sehr gut abbilden, weshalb die Bezeichnung Gesetzmäßigkeit naheliegt.

Informetrie wurde von Otto Nacke 1979 erstmals als Begriff erwähnt. Nacke definierte die Informetrie absichtlich sehr generell, als Oberbegriff zu den damals bereits bekannten Begriffen, wie „Librametry“, „Bibliometry“ oder „Scientometrics“ und erwähnte dabei explizit die Gesetzmäßigkeiten des Informationswesens.

„Informetrie ist die Lehre von der Anwendung mathematischer Methoden auf die Sachverhalte des Informationswesens zur Beschreibung und Analyse ihrer Phänomene, zum Auffinden ihrer Gesetze und zur Unterstützung ihrer Entscheidungen.“ (Nacke, 1979)

Leo Egghe, der jüngst eine Zeitschrift für Informetrie gegründet hat, kommt zu einer ähnlichen Definition. Informetrie ist auch bei ihm ein Ober- und Sammelbegriff für alle bislang bekannten Metriken:

„... we will use the term “informetrics” as the broad term comprising all metrics studies related to information science, including bibliometrics (bibliographies, libraries, ...), scientometrics (science policy, citation analysis, research evaluation, ...), webometrics (metrics of the web, the Internet or other social networks such as citation or collaboration networks), ...“ (Egghe, 2005)

Die „informetrischen Gesetzmäßigkeiten“, insbesondere Lotka's Law (Lotka, 1926), Bradford's Law of Scattering (Bradford, 1934) und Zipf's Law (1949), spielen seit den 1960er Jahren eine große Rolle in der Informetrie (siehe dazu (Brookes, 1977, Buckland, 1972, Garfield, 1980)). Mehrere Reviews und andere Publikationen haben sich mit der Informetrie und z. B. den informetrischen Gesetzmäßigkeiten auseinandergesetzt (Bar-Ilan, 2008, Egghe und Rousseau, 1990, Lockett, 1989, Wilson, 1999).

“An important subfield of informetrics, of long standing and of considerable sophistication, is the study of the informetric laws. ... the principal properties of interest are author(s), the publication journal, and the text words themselves; but, ..., the laws apply also to other informetric properties.” (Wilson, 1999)

4.1 Einführung des BLS

Der folgende Abschnitt führt die Gesetzmäßigkeit des BLS kurz ein und gibt Beispiele zum besseren Verständnis. Bei der Darstellung sollen keine mathematischen Details der Verteilungen diskutiert, sondern lediglich die wichtigsten Eigenschaften der typischen Verteilungen zusammengetragen werden.

Dem BLS liegt zugrunde, das sich die Literatur zu einem beliebigen Fachgebiet bzw. -thema z. B. in einer Bibliographie, in Bereiche mit unterschiedlichen Dokumentenkonzentrationen unterteilen lässt. So besteht zwischen den Zeitschriften eines Forschungsthemas und den Artikeln in diesen Zeitschriften eine quantifizierbare Relation, die Bradford als erstes beschrieben hat. Dem Kernbereich mit hoher Konzentration der Literatur folgen Bereiche mit zunächst mittlerer und geringer Konzentration, die jeweils die gleiche Menge an Zeitschriftenartikeln beinhalten wie der Kernbereich. Bradford beschreibt diese Beobachtung mit:

“The whole range of periodicals thus acts as a family of successive generations of diminishing kinship, each generation being greater in number than the preceding, and each constituent of a generation inversely according to its degree of remoteness.” (Bradford, 1934)

Samuel Bradford hat das nach ihm benannte informetrische Gesetz wie folgt formuliert:

“... if scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several groups or zones containing the same number of articles as the nucleus, when the numbers of periodicals in the nucleus and succeeding zones will be as 1 : n : n² ...” ((Bradford, 1934) S. 116)

Die Besonderheit der Bradfordschen Beobachtung liegt vor allem in der Betonung der Rangfolge bzw. dem Ranking der Zeitschriften nach der Anzahl der Artikel zu dem untersuchten Thema. Das gleiche Verfahren wird bei den verwandten Gesetzmäßigkeiten Lotka's Law für Publikationen eines Autors sowie Worthäufigkeiten in Texten in Zipf's Law angewendet (siehe dazu (Price, 1976)). BLS weist der produktivsten Zeitschrift den Rank 1 zu, der zweitproduktivsten Zeitschrift den Rank 2 usw. Auf dem letzten Rank befinden sich die Zeitschriften, die die geringste Anzahl an Artikeln zu dem Thema produzieren, das heißt, i. d. R. mindestens einen Artikel beinhalten (siehe Rangfolge in Tabelle 11).

(Egghe und Rousseau, 1990) nennen diese spezifischen Verteilungen „rank-order distributions“. Es werden in der Literatur aber auch Begriffe wie „long tail distributions“, „extremely skewed“, „law of the vital few“ oder auch „schiefe Verteilung“ verwendet. Power Laws bzw. Potenzgesetze folgen der gleichen Verteilung (vgl. (Newman, 2005)).

„Die Schiefeit bibliometrischer Verteilungen ist Ausdruck des nichtlinearen Charakters der zugrunde liegenden stochastischen Prozesse und weist auf das Vorhandensein langreichweitiger Wechselwirkungen hin.“ (Bonitz und Scharnhorst, 2001)

Im Prinzip stellt das BLS lediglich eine Präzisierung der 80 : 20-Regel für den Bereich der Zeitschriftenliteratur dar. 80 % der Nachfrage an Literatur können mit etwa 20 % des Bestandes in einer Bibliothek abgedeckt werden (vgl. (Umstätter, 2005, Wagner-Döbler, 1997)). Anders formuliert: 80 % der Zeitschriftenartikel zu einem Thema finden sich in 20 % der Zeitschriften, die zu diesem Thema publizieren. BLS hat damit Parallelen zu Paretos Beobachtungen auf dem Gebiet der Einkommensverteilungen.

Folgendes Beispiel soll zur Illustration der Bradfordschen Gesetzmäßigkeit dienen. Angenommen, es werden 3.000 Artikel zu einem bestimmten Thema in insgesamt 800 Zeitschriften gefunden. Dann ergibt sich in der oben beschriebenen Reihung nach Bradford bei drei gleich großen Gruppen von jeweils 1.000 Artikeln ein Verhältnis von 42 : 158 : 600 Zeitschriften. Näherungsweise $1 : 3,8 : 3,8^2$

Der Core oder Kern der Zeitschriften besteht demnach aus insgesamt 42 Zeitschriften. Diese 42 Zeitschriften beinhalten die ersten 1.000 Artikel. In der zweiten Zone (Zone 2) befinden sich 158 Zeitschriften, die wiederum 1.000 Artikel beinhalten. Die 600 Zeitschriften aus der dritten Zone (Zone 3) gehören zu den Peripheriezeitschriften, die nur selten, aber wenn dann mindestens einen thematisch relevanten Artikel beinhalten. Die Konzentration der Zeitschriftenartikel in den Zeitschriften nimmt somit pro Zone konstant ab.

Die von Bradford eingeführte Konstante⁵⁷, in diesem Beispiel 3,8, dient als Multiplikator. Das obere Beispiel lässt sich damit auch folgendermaßen formulieren: $42 : 42 \cdot 3,8 : 42 \cdot 3,8^2$.

Zu den Zonen 1, 2 und 3 (im Zitat „a, b and c groups“) findet sich bei Bradford folgende Aussage:

“... a and b groups above are periodicals of scope obviously and a priori relevant to the subjects investigated. They are journals which would almost inevitably suggest themselves as being concerned with these subjects, among others. ... the c groups are very mixed, including some sources which, by their scope, we might expect to find more productive; but including also journals of very general scope in which we should hardly expect to find references. ... We can only draw the general conclusion that a large number of references are produced by sources which, a priori, are “unlikely”. (Bradford, 1934)

Die Einteilung der beobachteten Objekte (Zeitschriften) in drei Zonen bei Bradford ist eine pragmatische Entscheidung, die keinen Zusammenhang mit der eigentlichen Gesetzmäßigkeit hat. Obwohl Bradford nur drei Zonen bzw. Gruppen zugrundelegt, sind aufgrund der „Selbstähnlichkeit“ (self-similarity) dieser Verteilungen auch mehr Zonen

57 Die Konstante ist als Bradford-Konstante beschrieben. Zur Analyse dieser Bradford-Konstante (bei Bradford auch Multiplier genannt) siehe Rao, Ravichandra I. K. (1998): An analysis of Bradford multipliers and a model to explain law of scattering, *Scientometrics* 41 [1-2], Seite 93-100.

möglich. (Egghe und Rousseau, 1990) zeigen dies an einem Beispiel (siehe Abbildung 4-1).

Die umfangreiche Literatur zu BLS lässt sich in unterschiedliche Gruppen einteilen: a) die Bestätigung der Bradford-Gesetzmäßigkeit in unterschiedlichen Fachgebieten und Disziplinen, b) die Diskussion über die mathematisch geeignete und korrekte Formulierung der Gesetzmäßigkeit sowie die Beziehung von BLS zu anderen informatrischen Gesetzmäßigkeiten und c) die Diskussion der theoretischen und praktischen Implikationen des Gesetzes.

In dieser Arbeit sollen insbesondere die praktischen Implikationen des BLS untersucht werden.

4.2 Bradfords Daten

Bradford selber hat für die beobachtete Gesetzmäßigkeit kein mathematisches Modell entwickelt, sondern hat das Gesetz lediglich verbal (siehe oben) und graphisch beschrieben.

Bradford fand in seiner ersten Arbeit (1934) folgende Verteilung von Zeitschriften zu den beiden Themen „Applied geophysics“ und „Lubrication“.

Die Analyse einer Bibliographie zu „Applied geophysics“ ergab folgende Verteilung in drei Gruppen von Zeitschriften:

- Die ersten 9 Zeitschriften lieferten zusammen 429 Artikel,
- die nächsten 59 Zeitschriften lieferten zusammen 499 Artikel ($9 \cdot 5 = 45$),
- die letzten 258 Zeitschriften lieferten zusammen 404 Artikel ($9 \cdot 5 \cdot 5 = 225$).

Analog zum oberen Beispiel lässt sich die Verteilung näherungsweise formulieren: $9 : 9 \cdot 5 : 9 \cdot 5^2$. Die Konstante nimmt hier den Wert 5 an.

Die Analyse einer Bibliographie zum Thema „Lubrication“ hat folgende Verteilung in drei Gruppen erbracht (siehe dazu auch Tabelle 11):

- Die ersten 8 Zeitschriften lieferten zusammen 110 Artikel (8 Zeitschriften),
- die nächsten 29 Zeitschriften lieferten zusammen 133 Artikel ($8 \cdot 3 = 24$ Zeitschriften),
- die letzten 127 Zeitschriften lieferten zusammen 152 Artikel ($8 \cdot 3 \cdot 3 = 72$ Zeitschriften).

Analog zum oberen Beispiel lässt sich die Verteilung näherungsweise formulieren: $8 : 8 \cdot 3 : 8 \cdot 3^2$. Die Konstante nimmt hier den Wert 3 an.

Die oberen Beispiele der Bradford-Daten zeigen, dass es sich dabei jeweils nur um Annäherungen handelt und eine exakte Berechnung der Größen der Zeitschriftengruppen und der Konstanten nicht vorgenommen werden kann. Die Konzentration der Zeitschriften kann zwischen den Gruppen durchaus variieren und die Konstanten passen daher nicht immer für alle Zonen gleich gut (siehe Beispiel „Lubrication“). (Egghe und Rousseau, 1990) zeigen das recht deutlich (siehe z. B. Abbildung 4-1).

Charakteristisch und allen Bradford-Verteilungen gemeinsam ist vielmehr die auftretende Streuung (Ungleichverteilung) bzgl. der Artikelproduktivität einzelner Zeitschriften zwischen den Zonen, die sich annähernd wie ein Power Law verhält.

„ ... all of which tend to demonstrate that certain systems, i.e., populations of sources and their activities recorded in some common manner, produce highly skewed distributions where most of the items are concentrated in a small population of sources, while the remaining items are spread out over the rest of the population ...” (Oluic-Vukovic, 1997) S. 833)

| | # journals | # articles | k |
|-----------------------|--|------------|------|
| 1 st group | $r_0 = 2.95 \approx 3$ | 55 | - |
| 2 nd group | $r_0 k = 4.98 \approx 5$ | 55 | 1.67 |
| 3 rd group | $r_0 k^2 = 8.42 \approx 8$ | 56 | 1.60 |
| 4 th group | $r_0 k^3 = 14.23 \approx 14$ | 56 | 1.75 |
| 5 th group | $r_0 k^4 = 24.05 \approx 24$ | 55 | 1.71 |
| 6 th group | $r_0 k^5 = 40.64 \approx 41$ | 49 | 1.71 |
| 7 th group | $r_0 k^6 = 68.68 \approx 69$ which is exactly the last existing rank | 69 | 1.68 |

Abbildung 4-1: Einteilung der Zeitschriften und Artikel zum Thema „Lubrication“ in 7 Gruppen (aus (Egghe und Rousseau, 1990) S. 344)

(Egghe und Rousseau, 1990) zeigen mit dem Beispiel, dass eine Bradford-Verteilung nicht unbedingt in die bei Bradford ursprünglich beschriebenen drei Gruppen eingeteilt werden muss, sondern auch in mehrere Zeitschriftengruppen aufgeteilt werden kann. Abbildung 4-1 zeigt die gleiche Verteilung der Zeitschriften zum Thema „Lubrication“ in sieben Gruppen.

Tabelle 11: Originaldaten Bradford (1934), Thema: „Lubrication“. Kumulierte Anzahl der Zeitschriften, Zeitschriften und die Anzahl der Artikel in diesen Zeitschriften, kumulierte Anzahl der Artikel.

| Zeitschriften kumuliert | Zeitschriften | Artikel | Artikel kumuliert |
|-------------------------|---------------|---------|-------------------|
| 1 | 1 | 22 | 22 |
| 2 | 1 | 18 | 40 |
| 3 | 1 | 15 | 55 |
| 5 | 2 | 13 | 81 |
| 7 | 2 | 10 | 101 |
| 8 | 1 | 9 | 110 |
| 11 | 3 | 8 | 134 |
| 14 | 3 | 7 | 155 |
| 15 | 1 | 6 | 161 |
| 22 | 7 | 5 | 196 |
| 24 | 2 | 4 | 204 |
| 37 | 13 | 3 | 243 |
| 62 | 25 | 2 | 293 |
| 164 | 102 | 1 | 395 |

Tabelle 11 präsentiert die Originaldaten von 1934 zum Thema „Lubrication“. Die erste Spalte („Zeitschriften kumuliert“) enthält die kumulierte Anzahl der Zeitschriften (Spalte 2). Die beiden folgenden Spalten („Zeitschriften“, „Artikel“) präsentieren die Anzahl der Zeitschriften und die jeweilige Anzahl der Artikel in diesen Zeitschriften. Die Zeitschrift mit der höchsten Artikelkonzentration (Rank 1) enthält demnach 22 Artikel zu dem Thema, die zweite Zeitschrift 18 Artikel (Rank 2) usw. Die letzte Spalte („Artikel kumuliert“) zeigt die kumulierte Anzahl der Zeitschriftenartikel in der Rangfolge der Zeitschriften.

Neben der verbalen Beschreibung der Gesetzmäßigkeit hat Bradford die Verteilungen der Artikel auf Zeitschriften graphisch visualisiert (siehe dazu Abbildung 4-2). Die graphische Darstellung wird folgendermaßen beschrieben:

“Along the x-axis Bradford ranked the journals 1, 2, 3 ... n in decreasing order of productivity of papers relevant to the given subject on a logarithmic scale. Along the y-axis he marked cumulative totals of papers, R(n). When R(n) is plotted against log n, the resulting graph begins with a rising curve which at a critical point P₁ runs into a straight line. The points X₁, X₂ and X₃ show the logarithm of the numbers of core journals (X₁), journals in Zone 2 (X₂) and journals in Zone 3 (X₃).” ((von Ungern-Sternberg, 2000) S. 163)

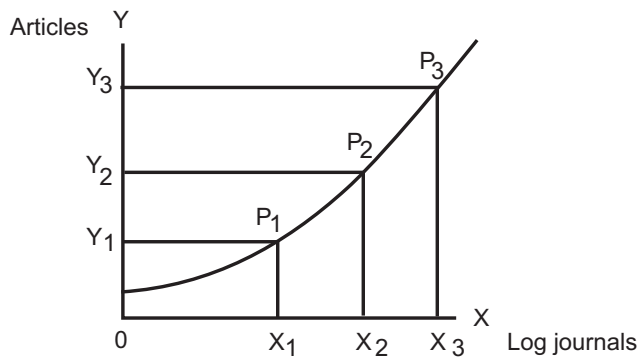


Abbildung 4-2: Graphische Darstellung der Bradford-typischen Verteilung (aus (von Ungern-Sternberg, 2000) S. 164)

Die Nachstellung der Originaldaten des Themas „Lubrication“ (siehe Tabelle 11) ergibt folgendes Bild:

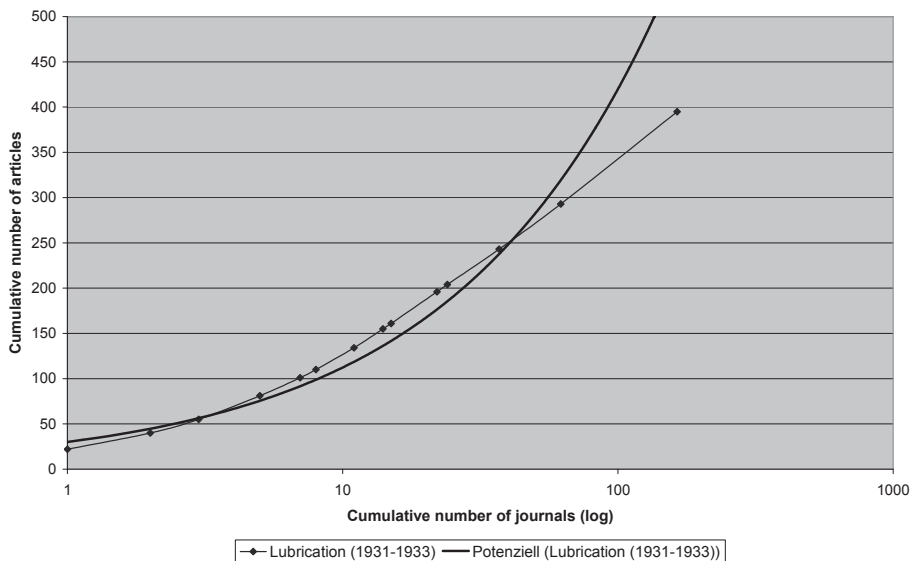


Abbildung 4-3: Originaldaten des Bradford-Themas „Lubrication“

Abbildung 4-2 zeigt charakteristische Bereiche der Bradford-Verteilung. Zum einen den Anstieg der Kurve am Anfang (Kernzeitschriften), zum anderen den geraden Verlauf im mittleren Bereich der Kurve. Die Trendlinie in Abbildung 4-3 folgt einem Power Law und zeigt die Ähnlichkeit der Bradfordschen Beobachtung zu diesem Typ Verteilung. Der für Bradford-Verteilungen typische Abfall der Kurve, der sog. „Groos droop“ (beschrieben z. B. bei (Brookes, 1968, Eto, 1988, Qiu und Tague, 1990)) im letzten Drittel der Kurve,

wird in der oberen Abbildung 4-3 nicht deutlich. Abbildung 4-4 visualisiert den Groos droop einer Bradford-Verteilung graphisch. Aus der J-förmigen Verteilung scheint dabei eine S-förmige zu werden.

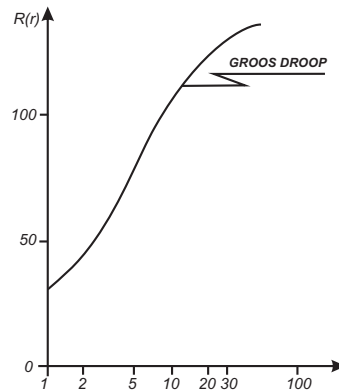


Abbildung 4-4: Visualisierung des Groos droop (aus (Egghe und Rousseau, 1990))

Der Groos droop beschreibt einen Sättigungseffekt bei Bradford-Verteilungen, der sich z. B. durch die Unvollständigkeit von Bibliographien erklären lässt. So wird die Recherche zu einem Thema in einer einzelnen Datenbank alleine i. d. R. nicht zu einer vollständigen Bibliographie des Themas führen (vgl. (Mayr und Umstätter, 2007)). Die Ausweitung der Recherche auf mehrere fachlich relevante Datenbanken, wie sie die Heterogenitätsbehandlung angestrebt (vgl. Abschnitt 3.2.1), wird mit großer Wahrscheinlichkeit weitere relevante Dokumente ergeben, die zu einer Minderung des Groos droop führen können. Der Groos droop hängt damit stark mit der gewählten Recherchestrategie zusammen (vgl. (Umstätter, 2005)).

“The Groos droop ... indicating that the search was not completed, ... and that the Groos droop can never be fully corrected.” ((Brookes, 1968) S. 252)

Letztlich zeigt der Groos droop nichts anders als eine typische und z. T. erklärbare Abweichung der empirischen Gesetzmäßigkeit. Da es sich beim BLS nicht um ein exaktes Gesetz handelt, sind Abweichungen dieser Art nicht weiter problematisch.

4.3 Einsatzmöglichkeiten

Bisher sind in der Literatur unterschiedliche Einsatzmöglichkeiten des BLS beschrieben, von denen hier nur einige genannt werden sollen. BLS dient dabei

- zum Aufbau, Management und zur Kontrolle von Kollektionen ((Bensman, 2001, Brookes, 1968, Brookes, 1969, Lancaster et al., 1991, Wilson, 1998)),
- zur Auswahl von Zeitschriften für eine Fachdatenbank (siehe (Triolo und Bao, 1993, von Ungern-Sternberg, 2000) und z. B. bei Thomson Scientific),
- zur Messung der Abdeckung von Bibliographien ((Goffman und Warren, 1969, Heine, 1998)) und

- zur Rechercheunterstützung (vor allem (Bates, 2002, Ingwersen, 1984, White, 1981)).

(Mayr und Umstätter, 2008) haben eine weitere Einsatzmöglichkeit beschrieben:

„Basierend auf der Erkenntnis, dass sich parallel zur Entstehung von 300 neuen Zeitschriften ein neues Literaturdokumentationssystem bildet, wie es D.J. de Solla Price beschrieben hat, entsteht bei 200 bis 300 Aufsätzen zu einem neuen Spezialthema pro Jahr eine neue Zeitschrift (Umstätter und Rehm, 1984). Bevor die Herausgeber einer neuen wissenschaftlichen Zeitschrift dieses Erfordernis für notwendig halten, brauchen sie nicht nur durchschnittlich hundert publikationsfähige Aufsätze pro Jahr, von etwa hundert Wissenschaftlern (Umstätter, 2003), um das eigene Core Journal zu füllen, sie brauchen vielmehr das Interesse von einigen hundert Abnehmern. (...) Im Sinne Bradford's Law of Scattering ist es möglich, die entsprechenden Werte aus den Kernzeitschriften heraus zu extrapolieren. Im Prinzip kann man abschätzen, dass die Gründung einer neuen Zeitschrift angezeigt erscheint, wenn die Zahl der wissenschaftlichen Publikationen zu einem Thema etwa die Zahl von 200 bis 300 pro Jahr erreicht. Da das Bradfordsche Gesetz besagt, dass im Prinzip jede der weltweit erscheinenden Zeitschriften für die Publikation dieses Themas in Frage kommt, auch wenn die Wahrscheinlichkeit, dem Potenzgesetz folgend, asymptotisch gegen Null geht, je fachfremder eine Zeitschrift wird.“ (Mayr und Umstätter, 2008)

Bradford selber hat in seinem Buch „Documentation“ (Bradford, 1948) ein Kapitel „The Documentary Chaos“ geschrieben und in diesem sein Verteilungsgesetz als Argument dafür genommen, Bibliographien künftig nach Zeitschriften auszuwerten und nicht nur nach rein fachlichen Gesichtspunkten aufzubauen. Bradford beschreibt damit sehr früh die Verfahren, mit denen noch heute Fachdatenbanken und Bibliographien aufgebaut werden. Bereits 1934 formuliert Bradford:

„It follows that the only way to glean all the articles on these subjects would be to scrutinise continually several thousands of journals, the bulk of which would yield only occasional references or none at all. And because this work is altogether impracticable, a large portion of the total number of references on a given subject is regularly missed by the abstracting journals. ... In methods of abstracting and indexing, a radical change is needed. Periodical literature must be abstracted by source, and not by subject, as hitherto.“ ((Bradford, 1934) S. 179-180)

Die Potenziale und Einsatzmöglichkeiten des BLS zum Bestandsmanagement bzw. dem Aufbau von Kollektionen finden sich auch bei (Brookes, 1968), einem frühen Verfechter dieser Gesetzmäßigkeit:

“Up to the present, information systems have tended to evolve, documentalists using their experience and making ad hoc adjustments to satisfy the demands for service made on the systems. But the discovery and formulation of statistical laws of wide generality would enable documentalists to design information services and to answer, a priori, such questions as: What would be the costs? ... What is the optimum distribution? How can a given collection best be subdivided into collections ...? ... “ (Brookes, 1968)

Insbesondere die Einsatzmöglichkeiten des BLS für die Informationssuche bei Bates (Bates, 1989, Bates, 2002) sind relevant für diese Arbeit und sollen hier weiter ausgeführt werden (siehe dazu auch (Mayr, 2006)):

“... the key point is that the distribution tells us that information is neither randomly scattered, nor handily concentrated in a single location. Instead, information scatters in a characteristic pattern, a pattern that should have obvious implications for how that information can most successfully and efficiently be sought.” (Bates, 2002)

Es lassen sich unterschiedliche Szenarien denken, in denen ein Nutzer die Suche nur auf die Kernjournale eines Fachgebiets einschränken will (z.B. Überblicksrecherche). Auf die Frage „Wie viele Datenbanken sind notwendig, um ein multidisziplinäres Forschungsthema zu recherchieren?“ kann eine Analyse nach Bradford beispielsweise die Datenbanken hervorheben, die die meisten Dokumente zu diesem Thema liefern (siehe dazu auch (Tenopir, 1982, Wilson, 1999)).

Im Sinne der Rechercheunterstützung kann davon ausgegangen werden, dass das Verteilungsmuster des BLS (Kernzone mit hoher Artikelhäufigkeit) für den Benutzer (vermutlich insbesondere ungeübte Nutzer und Anfänger) positive Auswirkungen beim Information Retrieval haben kann (vgl. (Ding et al., 2000, Wolfram, 2003)). Die Such- und Browsingtechniken des Berry Picking-Modells (Bates, 1989, Bates, 2002), aber auch andere Information Seeking-Modelle (Fisher et al., 2005, Marchionini, 1995), scheinen als Ansatzpunkt für das BLS durchaus geeignet.

„... having an understanding of the dynamics of literature distributions may enable us as searchers to make better and more sophisticated decisions about how we want to search and where. It may also prove valuable to test the general conjecture made here, that areas with high numbers of topically relevant materials (relative to all materials in the area) are best searched by browsing, areas with middling numbers of topically relevant items are best searched by directed searching on information-organizational structures, and areas with very sparse (“needle in a haystack”) numbers of relevant items are best searched by using links.” (Bates, 2002)

4.4 Bradfordizing-Effekt

Das Verfahren „Bradfordizing“ basiert auf der informatrischen Gesetzmäßigkeit des BLS. Das Verfahren Bradfordizing wurde erstmals in (White, 1981) zur Umorganisation von Suchergebnissen in Fachdatenbanken vorgeschlagen. Es sortiert bzw. rankt eine Dokumentmenge nach Kernzeitschriften. Die Zeitschriften zu einer Suchanfrage werden gemäß ihrer Häufigkeit (Anzahl der Zeitschriftenaufsätze) gelistet. Aufsätze der Zeitschriften aus Zone 1 werden beim Bradfordizing vor die Zeitschriftenartikel geordnet, die nur durchschnittlich (Zone 2) oder selten (Zone 3) gelistet werden. Das Verfahren stellt somit sicher, dass die Aufsätze zentraler Zeitschriften, oder auch Kernzeitschriften eines Gebiets, vor die Zeitschriften sortiert werden, die nicht zu den Kernpublikationen einer Disziplin oder eines Fachgebiets gehören. Bradfordizing ist insofern ein innovatives Re-Ranking-Verfahren, als es für eine beliebige Suchanfrage über eine nachgeordnete Umorganisation des Suchergebnisses zuverlässig (hohe Robustheit dieser Häufigkeitsverteilung) die zentralen Publikationen für eine übergreifende Suche in der Ergebnisliste nach vorne sortiert.

„... es [das Bradfordizing, Anmerkung des Autors] erleichtert nicht nur das Auffinden von Aufsätzen und hilft Anfängern rasch zu erkennen, welche Zeitschriften für sie besonders wichtig sind, es erhöht auch die Wahrscheinlichkeit, gerade in diesen Quellen noch weitere wichtige Publikationen zu finden ...“ (Umstätter, 2005)

Der Ablauf einer Bradfordizing-Analyse lässt sich folgendermaßen zusammenfassen (siehe dazu auch Abbildung 4-5):

- Identifikation⁵⁸ von allen Dokumenten (Artikeln), die in einem Forschungsgebiet/Subject publiziert wurden, durch retrospektive Suche in einem ausreichend großen Zeitraum.
- Auflistung der Quellen (Zeitschriften), in denen die Artikel publiziert wurden, in der Reihenfolge der Häufigkeit der Quelle. Das Ergebnis ist eine gerankte Liste der Zeitschriften (sog. „bradfordized list“).
- Einteilung der Zeitschriftenliste in Gruppen/Zonen (bei Beibehaltung der oberen Reihenfolge), so dass die Anzahl der Artikel pro Gruppe/Zone in etwa gleich ist.

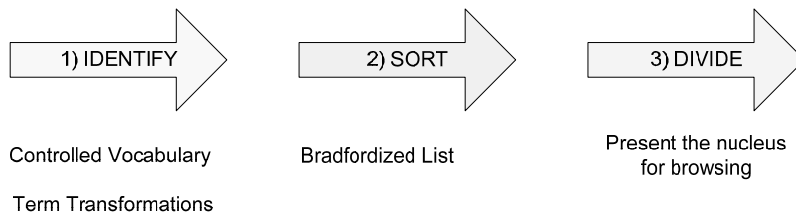


Abbildung 4-5: Schritte der Bradfordizing-Analyse

Eine detaillierte Beschreibung der einzelnen in dieser Arbeit angewendeten Analyse-schritte des Verfahrens Bradfordizing findet sich in Abschnitt 5.1.

Ziel der Anwendung von Bradfordizing in dieser Untersuchung ist es, den Re-Ranking-Ansatz („Ranking nach Kernzeitschriften“ oder Bradfordizing) für Zeitschriftenartikel und Monographien zu etablieren und in einem realen Suchszenario zu evaluieren (siehe dazu (Mayr, 2008)). Die Hauptidee der Anwendung von Bradfordizing ist es, ein beliebiges thematisch spezifiziertes Suchergebnis durch das einfach anwendbare Verfahren umzuorganisieren (Re-Ranking) und den Effekt dieses Re-Rankings zu evaluieren. Ziel ist es, nach dem Re-Ranking mehr relevante Dokumente für ein Topic im ersten Drittel der Ergebnismenge zu listen. Dieser Zugewinn an relevanten Dokumenten im Kern soll als Bradfordizing-Effekt bezeichnet werden.

Die Fokussierung auf Bradfordizing ist vielversprechend, weil dieses Verfahren per se universell und disziplinübergreifend angelegt ist und zuverlässig sowohl innerhalb einer Datenbank, einer Domäne (mehreren Datenbanken zu einem Fachgebiet) als auch

58 Zur Problematik der Operationalisierung einer Bradford-Analyse siehe Nicolaisen, Jeppe und Hjørland, Birger (2007): Practical potentials of Bradford's law: A critical examination of the received view, *Journal of Documentation* 63 [3], Seite 359-377.

zwischen Domänen beobachtet werden kann. In der Literatur finden sich viele Beispiele für die Bestätigung der Bradfordschen Beobachtung.

Sehr abstrakt ausgedrückt soll Bradfordizing durch die Identifikation von Kernbereichen als ein Kompensationsverfahren für erweiterte Suchräume (mehrere Datenbanken in einer Suchanfrage) eingesetzt werden. Der Heterogenität und der Vielzahl der Datenbanken wird als Ausgleich ein quasi vereinheitlichendes Prinzip entgegengesetzt, das Kernbereiche identifiziert, die über Datenbanken und Domänen hinweg existieren.

Bradfordizing liefert folgende unmittelbare Mehrwerte für den Nutzer:

- eine alternative Sicht auf Suchergebnisse, die nach Kernzeitschriften sortiert sind,
- eine alternative Sicht auf Publikationsquellen innerhalb eines Suchraums, die intuitiv näher (vgl. (Bates, 2002)) am Forschungsprozess liegt als statistische Verfahren (z. B. best match) oder traditionelle boolesche Verfahren (exact match),
- eine vermutlich höhere fachliche Relevanz (topicality) der Dokumente nach dem Re-Ranking.

Aufgrund der Robustheit und Allgemeingültigkeit des BLS für die Verteilung von Forschungsliteratur ist davon auszugehen, dass das Bradfordizing gerade in förderierten und interdisziplinären Suchumgebungen mit unterschiedlichen Informationsbeständen praktisch operabel und gewinnbringend ist (vgl. Ansatz bei (Chernov et al., 2006)).

Nachdem in diesem Kapitel Bradford's Law of Scattering und die Anwendung von Bradfordizing beschrieben wurden, soll im folgenden Kapitel auf das Untersuchungsdesign dieser Arbeit eingegangen werden.

5 Material und Methode

In den folgenden Abschnitten werden die einzelnen Information Retrieval (IR)-Untersuchungen dieser Studie sowie deren Evaluationsdesign beschrieben. Für die Tests wurden Daten aus zwei umfangreichen Evaluationsprojekten verwendet: zum einen Fragestellungen und Bewertungsdaten der jährlich stattfindenden internationalen CLEF-Initiative⁵⁹, zum anderen Daten, die im Rahmen des GESIS-Projekts KoMoHe im Jahr 2007 entstanden sind.

Allgemeine Ziele der CLEF-Initiative sind:

- Testen von IR-Systemen
- Bewertung von IR-Systemen (Fokus auf europäische Sprachen)
- Schaffen von Testsuiten wieder verwendbarer Daten (benchmarking)

„Generell geht es darum, ein offenes Diskussionsforum für den Austausch von Forschungserfahrungen und -ideen zu schaffen und die Kommunikation zwischen Wissenschaft und Wirtschaft im Bereich CLIR zu fördern. ... Spezielles Augenmerk wird auf die Evaluierungsmethoden und Analyseverfahren für CLIR gerichtet.“ ((Kluck, 2002) S. 83)

Die beiden Evaluationsprojekte eignen sich besonders für diese Studie, da es sich bei den getesteten Dokumenten größtenteils um Zeitschriftenartikel und Monographien handelt, die bezüglich konkreter Fragestellungen bewertet wurden. Für jede Fragestellung (im Kontext Information Retrieval Topic genannt) stehen jeweils ein Datensatz mit gefundenen Dokumenten und die entsprechenden binären Bewertungsdaten (relevant oder nichtrelevant) zur weiteren Auswertung zur Verfügung. Die Dokumente bzw. die Metadaten der Dokumente (vgl. Abbildung 5-2) stammen alle aus wissenschaftlichen Fachdatenbanken der Fächer Sozialwissenschaften, Politikwissenschaft, Wirtschaftswissenschaften, Psychologie und Medizin (siehe Übersicht in Tabelle 16).

Die Besonderheit der CLEF-Evaluationsdaten ist, dass insgesamt 125 vollständig bewertete Topics der Jahre 2003 - 2007 vorlagen und genutzt werden konnten. Die Studie kann das in dieser Arbeit entwickelte Verfahren Bradfordizing (siehe Abschnitt 4.4) somit auf der Basis einer großen Menge von unterschiedlichen Fragestellungen evaluieren. Einschränkend muss zu den CLEF-Daten bemerkt werden, dass die Analysen nur Aussagen für Dokumente einer einzelnen Datenbank, der sozialwissenschaftlichen Datenbank SOLIS, ermöglichen.

Die Besonderheit der KoMoHe-Daten ist die Anzahl und Heterogenität der untersuchten Datenbanken. Die KoMoHe-Evaluation hat Dokumente aus 7 Datenbanken einbezogen, die alle mit spezifischen kontrollierten Vokabularen erschlossen werden und aus unterschiedlichen Disziplinen stammen. Zudem wurden zur Gewinnung zusätzlicher Dokumente intellektuelle Crosskonkordanzen eingesetzt (siehe Abschnitt 3.2.2).

In diesem Kapitel werden das angewendete Verfahren Bradfordizing, die untersuchten Testkorpora und Topics sowie die Beschreibung der IR-Tests in den jeweiligen Evalua-

59 <http://clef.iei.pi.cnr.it/>

tionsprojekten beschrieben. Zum Abschluss des Kapitels wird ein qualitatives Interview zu den Kernzeitschriften der Topics innerhalb der CLEF-Evaluation 2007 durchgeführt.

5.1 Bradfordizing

Bradfordizing ist ein Sortiervverfahren, das auf unterschiedliche Dokumenttypen in wissenschaftlichen Fachdatenbanken angewendet und auf Bradford's Law of Scattering zurückgeführt werden kann (siehe dazu ausführlicher Kapitel 4). Howard White hat das Verfahren Bradfordizing 1981 in einem Aufsatz in der Zeitschrift *Online* vorgestellt. White beschreibt das Verfahren folgendermaßen:

“... That is sorting hits (1) by the journal in which they appear, and then sorting these journals not alphabetically by title but (2) numerically, high to low, by number of hits each journal contains. In effect, this two-step sorting ranks the search output in the classic Bradford manner, so that the most productive, in terms of its yield of hits, is placed first; the second-most productive journal is second; and so on, down through the last rank of journals yielding only one hit apiece.” ((White, 1981) S. 47)

In dieser Studie haben wir Bradfordizing auf Dokumentenpools aus unterschiedlichen Fachdatenbanken aus den Bereichen Sozial- und Politikwissenschaften, Wirtschaftswissenschaften, Psychologie und Medizin zu verschiedenen Fragestellungen angewendet. Die Topics bzw. die Dokumentenpools zu diesen Topics dienen jeweils als Grundlage für die Bradfordizing-Analyse. Die Bradfordizing-Analyse unterteilt sich in folgende Schritte:

1. Generierung bzw. Zusammenstellung der Dokumente zu einem Topic,
2. Filterung der Dokumente nach Dokumenttyp,
3. Häufigkeitsanalyse (Pivot-Analyse) nach Dokumenttyp.

Ein Dokumentenpool mit n (n ist in der Regel >100) Dokumenten wird der Analyse zugrunde gelegt. In einem ersten Schritt werden die Dokumente zu einem Topic aus der Gesamtmenge isoliert, die mit dem Verfahren Bradfordizing umorganisiert werden sollen. In dieser Studie wird Bradfordizing auf die beiden Dokumenttypen Zeitschriftenartikel und Monographien angewendet. Nach dem zweiten Schritt liegen somit Zeitschriftenartikel oder Monographien zu einem Topic vor. Da die Zeitschriftenartikel im erweiterten GIRT-Format (siehe Abbildung 5-7) i. d. R. eine identifizierende Nummer der Zeitschrift (ISSN-Nummer) tragen, kann die anschließende Häufigkeitsanalyse (das eigentliche Bradfordizing) auf Basis der ISSN (International Standard Serial Number) erfolgen.

Das gleiche Verfahren kann auf die ISBN-Nummer bei den Monographien angewendet werden. Die ISBN (International Standard Book Number)⁶⁰ ist ein Identifier für Monographien und andere selbstständige Veröffentlichungen, in dem der Verlag, der die Publikation herausgibt, kodiert ist. Die Verlagsnummer ist eine ein- oder mehrstellige Ziffer, die eindeutig einem Verlag zugeordnet ist. Dieser Verlags-Code wurde für die Häufigkeitsanalyse (Bradfordizing) verwendet.

60 2007 wurde die ISBN auf 13 Stellen ausgeweitet. Soweit 13-stellige ISBNs in der Analyse vorkamen, wurde dies bei der Identifikation der Verlage berücksichtigt.

Folgendes vereinfachtes Beispiel mit unterschiedlich gehäuften ISSN-Nummern soll das Verfahren Bradfordizing bei Zeitschriftenartikeln exemplarisch illustrieren:

Tabelle 12: Liste mit individuellen ISSN-Nummern

| | ISSN |
|----|-----------|
| 1 | 0936-7780 |
| 2 | 0723-399X |
| 3 | 0341-7069 |
| 4 | 0933-1883 |
| 5 | 0933-1883 |
| 6 | 0933-1883 |
| 7 | 0028-3355 |
| 8 | 0030-9273 |
| 9 | 0171-3957 |
| 10 | 0171-3957 |

Aus einer beliebig zusammengesetzten Liste mit ISSN-Nummern zu einem Topic (siehe Tabelle 12) wird nach dem Bradfordizing eine nach Häufigkeit sortierte Liste dieser ISSN-Nummern (siehe Tabelle 13). Die ISSN, die am häufigsten in der Liste vorkommt, wird an die erste Stelle sortiert und die Häufigkeit wird in eine neue Spalte geschrieben. Eine weitere Spalte (rechte Spalte in Tabelle 13) kumuliert die Anzahl der Artikel pro Zeitschrift. Die kumulierten Werte der Artikel werden im Anschluss an das Bradfordizing benötigt, um die drei in etwa gleich großen Zonen zu bestimmen (analoges Verfahren für Monographien auf Basis der Verlags-Codes innerhalb der ISBN).

Tabelle 13: Liste der ISSN-Nummer nach Bradfordizing

| Rang | ISSN | Artikel | Artikel kumuliert |
|------|-----------|---------|-------------------|
| 1 | 0933-1883 | 3 | 3 |
| 2 | 0171-3957 | 2 | 5 |
| 3 | 0028-3355 | 1 | 6 |
| 4 | 0030-9273 | 1 | 7 |
| 5 | 0341-7069 | 1 | 8 |
| 6 | 0723-399X | 1 | 9 |
| 7 | 0936-7780 | 1 | 10 |

Die für diese Studie angeschlossene Zoneneinteilung auf Basis der kumulierten Werte in Core, Zone 2 und Zone 3 ist eine rein pragmatische Einteilung (in Anlehnung an Brad-

ford), die rechnerisch die Dokumentmenge pro Topic in etwa drei gleich große Teile teilt und die entsprechende Zonenzuteilung pro Zeitschrift bzw. Verlag annähert. Andere Einteilungen der Dokumente z. B. in vier oder mehr Zonen sind möglich, wurden in dieser Studie aber nicht operationalisiert.

Das folgende idealisierte Beispiel in Abbildung 5-1 visualisiert die drei Zonen nach einer Bradfordizing-Analyse. Die Dokumentmenge umfasst in diesem Beispiel 450 Zeitschriftenartikel zu einem Topic, die in dem Fall auf insgesamt 39 unterschiedliche Zeitschriften verteilt sind (150 Artikel in jeder der drei Zonen). Das Kumulieren der Artikelzahlen nach dem Bradfordizing hat in diesem Fall ergeben, dass die ersten drei Zeitschriften zusammen 150 Artikel ergeben, also das erste Drittel der Gesamtdokumentenzahl von 450 Dokumenten. Diese ersten 150 Dokumente in den drei Zeitschriften definieren damit den Nukleus oder Core für dieses Topic. Die drei Zeitschriften werden „Core Journals“ oder Kernzeitschriften genannt. Für das zweite Drittel bzw. Zone 2 werden 9 (3*3 Zeitschriften) weitere Zeitschriften und für das dritte Drittel bzw. Zone 3 werden 27 (3*3*3 Zeitschriften) Zeitschriften benötigt, um die Menge von jeweils 150 Zeitschriftenartikeln zu erreichen.

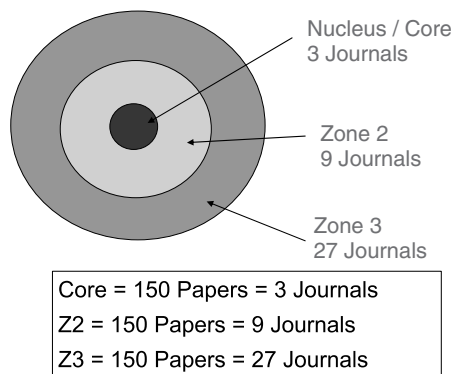


Abbildung 5-1: Ergebnis eines idealisierten Dokumentenpools für eine Fragestellung nach Bradfordizing. Einteilung in typische drei Bradford-Zonen (Core, Zone 2 und Zone 3). Die 450 Zeitschriftenartikel sind über 39 Zeitschriften verstreut.

Das Sortiervfahren Bradfordizing kann grundsätzlich für jedes beliebige Suchergebnis mit einer bestimmten Mindestmenge von Dokumenten (>100) angewendet werden. Es werden i. d. R. mehr als 100 Dokumente für das Bradfordizing benötigt, weil in kleineren Dokumentmengen nicht genügend Streuung vorliegt, um die Zonenbildung nach Bradford vernünftig einsetzen zu können.

Die Hauptidee der Anwendung von Bradfordizing in dieser Arbeit ist es, ein beliebiges thematisch spezifiziertes Suchergebnis durch ein einfach anwendbares Verfahren umzuorganisieren (Re-Ranking) und den Effekt dieses Re-Rankings zu evaluieren (siehe dazu Abschnitt 5.2.2).

Der folgende Abschnitt beschreibt die einzelnen Untersuchungen dieser Studie, die anhand der Bewertungsdaten der beiden Evaluationsprojekte KoMoHe und CLEF durchgeführt wurden.

5.2 Evaluation

Abbildung 5-2 visualisiert ein typisches Retrieval-Evaluationsszenario (siehe dazu auch (Binder et al., 2000, Harman und Voorhees, 2006, Womser-Hacker, 2004)). Die Organisatoren einer Retrieval-Konferenz (z. B. TREC oder CLEF) stellen zunächst die Testkollektion („Document set“) und Fragestellungen („Topics“) zur Verfügung. Die Teilnehmer wenden dann ihre individuellen Retrieval-Algorithmen auf die Fragestellungen und die Dokumente der Testkollektion an. Jedes teilnehmende System liefert eine gerankte Ergebnisliste pro Fragestellung an die Organisatoren, die daraus einen gemeinsamen Dokumentenpool erstellen (vgl. „document pools“). Die Dokumentenpools pro Topic werden im Anschluss von Juroren vollständig bewertet. Aus den intellektuellen Relevanzbewertungen lassen sich dann die bekannten Retrieval-Maße berechnen, mit denen die Leistungen der unterschiedlichen Retrieval-Algorithmen pro Topic verglichen werden können.

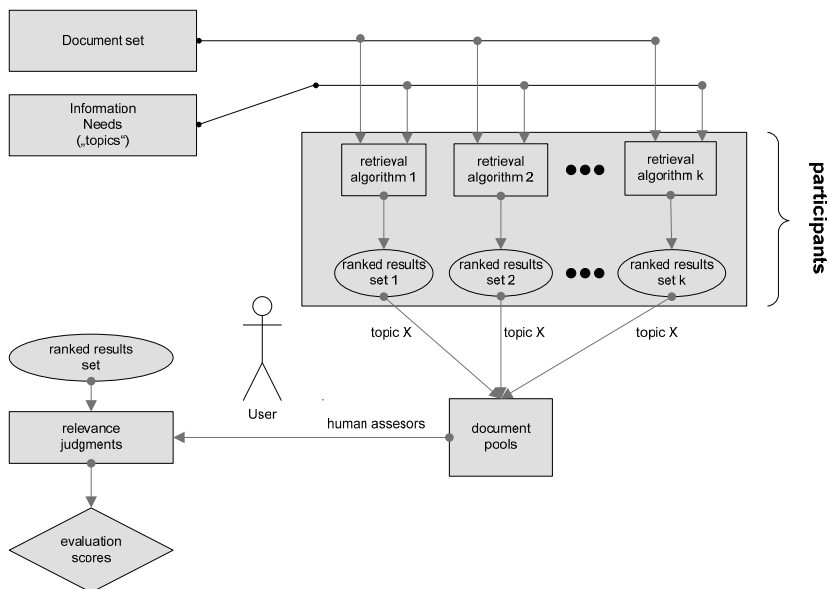


Abbildung 5-2: Typisches IR-Evaluationsszenario nach dem Verfahren TREC (nach (Voorhees, 2007))

Die Analyse der unterschiedlichen Topics in den Projekten CLEF (siehe (Petras et al., 2007)) und KoMoHe erfolgte nach dem Vorbild der klassischen Retrieval-Konferenzen (vgl. Abbildung 5-2).

Im Projekt KoMoHe wurden Dokumente aus 7 Datenbanken zu insgesamt 82 Fragestellungen bewertet und untersucht. Die Evaluation im Projekt KoMoHe teilt sich a) in die Untersuchung der Effektivität der Term-Transformationen (Crosskonkordanzen) und b) in die Anwendung des Re-Ranking-Verfahrens Bradfordizing auf die relevanzbewerten Daten der Untersuchung a). Die bewerteten Dokumente (vgl. ‚document pools‘ in Abbildung 5-2) **des ersten IR-Tests (Term-Transformationen) wurden für die anschließende Re-Ranking-Untersuchung weiter verwendet** (siehe dazu Abbildung 5-5). Mit den bewerteten Datensätzen der CLEF-Topics (125 Topics) wurde analog dazu verfahren. Die Analyse der Effektivität des Re-Rankings setzt jeweils auf den Dokumentpools einer konkreten Fragestellung auf.

Nachfolgend wird das Design der einzelnen Untersuchungen vorgestellt.

5.2.1 Evaluation der Term-Transformationen

Die Hauptidee der Nutzung der Crosskonkordanzen ist es, die Suchbegriffe einer Anfrage mit kontrollierten Termen in weitere kontrollierte Vokabulare zu übersetzen, um damit die Suche über unterschiedliche Datenbanken zu ermöglichen. Der Einsatz der Crosskonkordanzen soll damit den Suchraum erweitern, Ambiguitäten und Ungenauigkeiten in der Anfrageformulierung korrigieren und folglich mehr relevante Dokumente zu einer bestimmten Anfrage finden. Die Anwendung von Crosskonkordanzen bei der Suche stellt damit einen Mehrwert dar: Crosskonkordanzen können die Geschwindigkeit und Einfachheit des Suchprozesses beeinflussen. Die Suchraumerweiterung bzw. Erhöhung des Recalls für retrospektive Recherchen ist insbesondere für Untersuchungen, die auf Vollständigkeit zielen (z. B. bibliometrische Analysen), ein entscheidender Faktor und Mehrwert.

Obwohl die Notwendigkeit und der Nutzen von Terminologie-Mapping von der Community allgemein anerkannt werden und viele Mapping-Projekte durchgeführt wurden, wurde die Effektivität und Nützlichkeit der Projektergebnisse selten stringent evaluiert. Das Projekt KoMoHe konnte als erstes einen Teil der im Projekt entstandenen Terminologie-Mappings (Crosskonkordanzen) in einem standardisierten Information Retrieval-Test evaluieren.

Bereits bei (Marx, 2005) finden sich erste Ergebnisse der Evaluation von intellektuell und automatisch (statistisch) erstellten Term-Term-Relationen zwischen zwei unterschiedlichen Vokabularen. Marx kann zeigen, dass sich die Term-Precision und der Term-Recall zwischen Termen aus SWD und TheSoz für beide Transformationsarten (intellektuell durch Crosskonkordanzen und statistisch) gegenüber einer Nulltransformation (keine Transformation) sichtbar verbessern. (Zhang, 2006) kann dieses Ergebnis mit einem rough set-basierten Transferverfahren bestätigen. Die Evaluation von Term-Precision und Term-Recall bei Marx kann damit als Test im Sinne eines Schnelltests für die Evaluation von Term-Transformations-Verfahren angesehen werden. Aus diesem Grund wurde von Vivien Petras, Anne-Kathrin Walter und Philipp Mayr im Projekt KoMoHe ein spezifischer IR-Test konzipiert, der die Effektivität der intellektuell erstellten Crosskonkordanzen per Dokumentbewertungen evaluiert.

Der folgende IR-Test hat das Ziel, die Anwendung von Crosskonkordanzen bzw. Term-Transformationen⁶¹ in einem realen Suchszenario zu evaluieren⁶² (siehe dazu Fragestellung 1 in Kapitel 1). Im Unterschied zum typischen Evaluationsszenario (vgl. Abbildung 5-2) **werden nicht unterschiedliche Retrieval-Algorithmen miteinander verglichen, sondern vor allem die beiden Anfragetypen a) Anfrage ohne Term-Transformation und b) Anfrage mit Term-Transformation.** Als weitere Anfrageart wird c) die Freitext-Suche diesem IR-Test als Vergleichswert hinzugefügt. Die Ergebnismengen dieser drei Anfragetypen pro Topic werden wie im oberen Beispiel in einen gemeinsamen Dokumentpool zusammengeführt und anschließend bewertet.

Bei der Evaluation von Terminologie-Mappings ist der analytische Startpunkt entscheidend. Es ist zu klären, was untersucht wird: die Qualität der Mappings an sich oder die Qualität der anschließenden Suche? Die Qualität der Mappings ist eine Grundvoraussetzung für eine Verbesserung der Qualität der Suche. Daher wurden die Crosskonkordanzen im Projekt KoMoHe alle von den Fachexperten der Partnerinstitutionen geprüft. Die manuelle Erstellung und sorgfältige Prüfung gewährleistet, dass die Mappings sinnvoll, angemessen und von konsistenter Qualität sind.

Wenn die Qualität von Crosskonkordanzen evaluiert werden soll, kommen bei der Suche mehrere Faktoren zum Tragen: die Crosskonkordanzen an sich, aber auch die Inhalte der involvierten Datenbanken, deren Abdeckung oder Überlappung der Inhalte, die Suchoberfläche oder das eingesetzte Retrieval-Ranking. Das Ziel der Untersuchung war, die Auswirkungen der Crosskonkordanzen zu evaluieren und gleichzeitig die aktuellen Retrieval-Bedingungen (Oberfläche, Rankingmethode usw.) so weit wie möglich konstant zu halten. Vorteil der Untersuchung war, dass viele Crosskonkordanzen und ihre zugehörigen Datenbanken in-house getestet werden konnten, indem die Dokumente mit dem Open-Source Information Retrieval-System Solr⁶³ indexiert wurden. Es wurden die gleichen Processing- und Ranking-Module für alle Datenbanken verwendet. Für die Datenbanken, die nicht in-house verfügbar waren (z. B. Econis, Psynindex und Medline), haben wir die Datenbank-Hosts gebeten, gerankte Ergebnislisten für die vorbereiteten Anfragen zu liefern.

Für die meisten Datenbanken wurden die Term-Mappings in beide Richtungen getestet, ausgehend von Vokabular A nach B ($A \rightarrow B$) und umgekehrt von Vokabular B nach A ($B \rightarrow A$). Da die bilateralen Crosskonkordanzen unterschiedliche Suchen (unterschiedliche Anfragen, die vom jeweiligen Startvokabular abhängen) und Datenbanken involvieren, sind die Topics und demzufolge die Suchanfragen unabhängig voneinander zu betrachten (siehe dazu Übersicht der evaluierten Topics in Tabelle 14).

61 Im Folgenden werden die Begriffe Crosskonkordanz, Term-Transformation, Terminologie-Mapping und Mappings synonym verwendet.

62 Dieser Abschnitt basiert auf dem Artikel „Cross-concordances: terminology mapping and its effectiveness for information retrieval“ von Mayr, Philipp und Petras, Vivien (2008): Cross-concordances: terminology mapping and its effectiveness for information retrieval, IFLA World Library and Information Congress, Québec, Canada. URL: http://www.ifla.org/IV/ifla74/papers/129-Mayr_Petras-en.pdf.

63 <http://lucene.apache.org/solr/>

Eine Prämisse für die technische Implementation der Terminologie-Mappings sollte sein, dass die Mappings durch den Suchenden unkompliziert genutzt werden können. Die Mappings sollten das Sucherlebnis verbessern, ohne die Komplexität des Informationssystems für den Nutzer zu erhöhen. Die Nutzung der Crosskonkordanzen für die Evaluation folgte einem strikt automatischen Ansatz, d. h. es war kein manuelles Eingreifen in Form von intellektueller Anfrage-Reformulierung notwendig (siehe dazu aktuelle Implementation der Crosskonkordanzen im Suchsystem sowiport).

Zwei Information Retrieval-Tests wurden festgelegt, um die Effektivität der Crosskonkordanzen bei der Suche zu evaluieren:

Test A: Verbessert der Einsatz der Crosskonkordanzen die Suche gegenüber einer nicht-transformierten fachlichen (z. B. über ein kontrolliertes Vokabular) Suche?

In Test A wurde eine Anfrage in Terme eines kontrollierten Vokabulars (A) übersetzt und dann in den Schlagwort-Feldern einer bibliographischen Datenbank gesucht, die mit dem kontrollierten Vokabular (B) erschlossen ist (Szenario Controlled Terms (CT) in Abbildung 5-3). Die Suche wurde mit Hilfe der Crosskonkordanz $A \rightarrow B$ wiederholt, wobei die Suchterme des ursprünglichen kontrollierten Vokabulars mit Hilfe des Heterogenitätsservice (HTS, siehe Beschreibung in Abschnitt 3.2.5) in die kontrollierten Terme der Zieldatenbank übersetzt wurden (Szenario Term-Transformation (TT) in Abbildung 5-3). **Abbildung 5-3 zeigt eine graphische Repräsentation dieses Prozesses. Die Retrieval-Ergebnisse der beiden Szenarien CT (ohne Transformation) und TT (mit Transformation) wurden daraufhin verglichen.**

Wenn die Ergebnisse der beiden Suchen identisch sind, dann hat die Anwendung der Crosskonkordanzen für die Termtransformation keinen Effekt. Wenn sich die Suchergebnisse verschlechtern, dann haben die Crosskonkordanzen einen negativen Effekt. Wenn sich das Suchergebnis verbessert, was erwartet wird, dann hat der Einsatz der Crosskonkordanzen einen positiven Effekt auf die Suche.

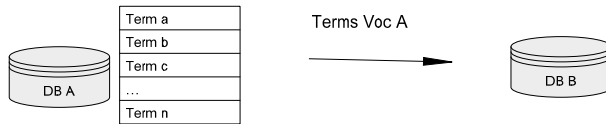
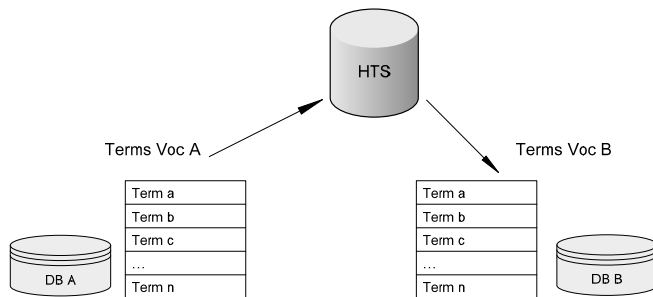
Scenario CT**Scenario TT**

Abbildung 5-3: Aufbau der Information Retrieval-Evaluation mit Crosskonkordanzen. Szenario CT = Suche mit kontrollierten Termen, Szenario TT = Suche mit Termtransformation, HTS = Heterogenitätsservice

Test B: Verbessert der Einsatz der Crosskonkordanzen die Freitext-Suche?

In Test B wurde die natürlichsprachige Originalanfrage in einem Szenario mit Freitext-Suche genutzt. Dieser Test wurde der Evaluation hinzugefügt, da die meisten Suchenden nicht die entsprechenden kontrollierten Vokabulare bei der Recherche verwenden. Bei der Freitext-Suche werden die Suchbegriffe nicht ausschließlich in den kontrollierten Vokabularfeldern gesucht, sondern zudem in den Titel- und Abstract-Feldern. In unserem Experiment wurde die Originalanfrage zuerst in den Freitext-Feldern der Datenbank gesucht (siehe Szenario FT in Abbildung 5-4). Anschließend wurden die Anfrageterme in den Crosskonkordanzen nach vorhandenen Term-Mappings geprüft. Die neuen Terme aus der Crosskonkordanz wurden über eine Boolesche OR-Verknüpfung zu der Originalanfrage hinzugefügt (siehe Szenario FT+TT in Abbildung 5-4).

Ein Beispiel illustriert die Unterschiede zwischen den beiden Tests A und B: eine natürlichsprachige Anfrage sucht nach Dokumenten zum Thema „family relations“ (dt. „Familienbeziehungen“). „Family relations“ ist bereits ein kontrollierter Term in dem Vokabular A und braucht deshalb nicht in einen kontrollierten Term übersetzt zu werden, er wird somit als Term für die erste Suche im Schlagwort-Feld der Datenbank B verwendet:

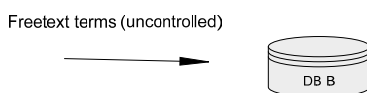
Test A CT: *Family relations*

Die Crosskonkordanz $A \rightarrow B$ bildet die Phrase „family relations“ aus Vokabular A auf die Termkombination „family“ AND „social relations“ in Vokabular B ab. Die zweite Suche in Datenbank B ist folglich:

Test A TT: *Family AND social relations*

Für den Test B wird die Originalanfrage in den Freitext-Feldern der Datenbank B gesucht (Titel, Abstract und kontrollierte Terme):

Scenario FT



Scenario FT+TT

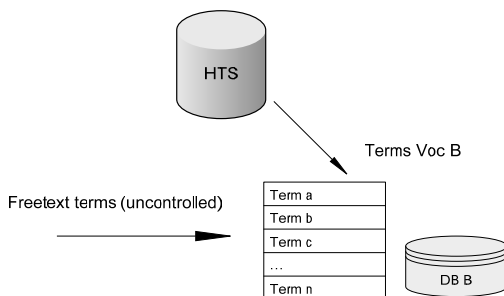


Abbildung 5-4: Aufbau der Information Retrieval-Evaluation von Freitext-Termen und Term-Transformationen

Test B FT: *Family relations*

Da die Anfrageterme in Vokabular A existieren, kann ein Term-Mapping für die Anfrage in den Crosskonkordanzen gefunden werden. Die Terme werden zu der Originalanfrage hinzugefügt und in den Freitext-Feldern der Datenbank B gesucht:

Test B FT+TT: *Family relations OR (Family AND social relations)*

Test A durchsucht nur die Schlagwort-Felder, wohingegen Test B zusätzlich weitere Felder (Titel und Abstract) durchsucht, in denen unter Umständen die Anfrageterme vorkommen. Test B ist generell ein schwächerer Test, da die transformierten kontrollierten Terme über eine OR-Verknüpfung der Anfrage hinzugefügt werden, aber nicht die Originalanfrage ersetzen wie in Test A. Da Test B auch die Felder aus Test A durchsucht, kann über Test B ausgesagt werden, dass er Test A subsumiert. In Test B ergeben sich jedoch

weniger Term-Erweiterungen, weil nicht alle Anfrageterme im originalen kontrollierten Vokabular auftauchen und somit weniger Term-Mappings gefunden werden.

Um zu gewährleisten, dass realistische Anfragen verwendet werden, wurde zur Anfrageerstellung die Hilfe der Produzenten oder Anbieter der getesteten Datenbanken erbeten. Die Anbieter wurden gebeten, 3 - 10 Anfragen (Durchschnitt: 6 - 7) aus ihrer täglichen Arbeit zur Verfügung zu stellen. Diese Anfragen wurden z. T. vereinfacht und anschließend in die kontrollierten Vokabulare der getesteten Datenbanken übersetzt. Die natürlichsprachigen Freitext-Anfragen enthielten ca. 1 - 3 Terme pro Anfrage, während die Booleschen Anfragen für die Schlagwortsuche ca. 2 - 6 Terme enthielten. Für alle Information Retrieval-Experimente wurden ausschließlich Äquivalenzrelationen der Term-Mappings verwendet. Es wurden aus allen Informationssystemen jeweils nur die ersten 1.000 gelisteten Dokumente einbezogen. Zum Schluss wurden die gepoolten Dokument-Ergebnismengen aus jedem Experiment bzgl. ihrer Relevanz zur Anfrage bewertet (siehe dazu Abschnitt 5.5).

Um den Effekt der Crosskonkordanzen zu evaluieren, wurden die klassischen Information Retrieval-Maße Recall und Precision auf Basis der Relevanzurteile der gefundenen Dokumente errechnet. Die folgenden Maßzahlen wurden analysiert:

- Retrieved: durchschnittliche Anzahl der gefundenen Dokumente (über alle Suchtypen),
- Relevant: durchschnittliche Anzahl der relevanten gefundenen Dokumente (über alle Suchtypen),
- Rel_ret: durchschnittliche Anzahl der relevanten gefundenen Dokumente eines bestimmten Suchtyps,
- Recall: Anteil der relevanten gefundenen Dokumente an allen relevanten Dokumenten (Durchschnitt über alle Anfragen eines Suchtyps),
- Precision: Anteil der relevanten gefundenen Dokumente an allen gefundenen Dokumenten (Durchschnitt über alle Anfragen eines Suchtyps),
- P10: Precision at 10 = Precision nach 10 gefundenen Dokumenten,
- P20: Precision at 20 = Precision nach 20 gefundenen Dokumenten.

P10 und P20 wurden berechnet, um ein realistisches Suchszenario zu repräsentieren, in dem ein Nutzer normalerweise nicht mehr als eine oder zwei Ergebnisseiten betrachtet. Für Retrievalsysteme, die nicht ranken, sondern die Ergebnisse nach Jahr oder Autoren sortieren, sind P10 und P20 weniger aussagekräftig.

Für beide Experimente wurden die Crosskonkordanzen nach Disziplin (intra- oder interdisziplinär) und nach Sprache (ein- oder mehrsprachig) aufgeteilt. Diese Aufteilung findet sich anschließend auch in der Ergebnisdarstellung. Intradisziplinäre Crosskonkordanzen im Test setzen sich hauptsächlich aus Vokabularen in den Sozialwissenschaften zusammen, da sich die meisten Crosskonkordanzen im Projekt KoMoHe in dieser Disziplin verorten. Die interdisziplinären Crosskonkordanzen bilden Vokabulare in den Disziplinen Wirtschaftswissenschaft, Medizin, Politikwissenschaft, Psychologie und den Sozialwissenschaften ab. Die einsprachigen Crosskonkordanzen schließen deutschsprachige Vokabulare ein; die zweisprachigen Crosskonkordanzen beinhalten deutsch- und

englischsprachige Vokabulare. Tabelle 14 gibt einen Überblick über die Anzahl der getesteten Crosskonkordanzen pro Experiment:

Tabelle 14: Anzahl der getesteten Crosskonkordanzen in der IR-Evaluation des Projekts KoMoHe

| Test A: kontrollierte Term-Suche (CT) | Schlagwort-Suche |
|---------------------------------------|------------------|
| intradisziplinäre Crosskonkordanzen | 5 (1 bilingual) |
| interdisziplinäre Crosskonkordanzen | 8 |
| Test B: Freitext-Suche (FT) | Freitext-Suche |
| intradisziplinäre Crosskonkordanzen | 6 (1 bilingual) |
| interdisziplinäre Crosskonkordanzen | 2 |

Der Hintergrund zur Aufteilung der Crosskonkordanzen in dieser Weise liegt in der Hypothese begründet, dass Crosskonkordanzen zwischen Vokabularen in der gleichen Disziplin (intradisziplinär) sich vermutlich mehr überlappen und mehr identische Terme enthalten und folglich einen geringeren Effekt auf die Retrieval-Ergebnisse haben als interdisziplinäre Crosskonkordanzen. Term-Mappings zwischen Vokabularen in unterschiedlichen natürlichen Sprachen (z. B. Englisch → Deutsch) oder zwischen unterschiedlichen Notationssystemen (z. B. Klassifikation → Thesaurus) werden vermutlich einen viel größeren Effekt haben, weil das Vorkommen von identischen Termen oder eine Überlappung unwahrscheinlicher sind.

Die Ergebnisse der Evaluation der Crosskonkordanzen bzw. Term-Transformationen finden sich in Abschnitt 7.1.

Im Anschluss an diese Tests wurden die relevanzbewerteten Dokumente aus den gepoolten Datensätzen pro Topic für das folgende Experiment weiterverwendet.

5.2.2 Evaluation des Re-Rankings

Die folgende Evaluation setzt auf den Evaluationsdaten des IR-Tests des vorherigen Abschnitts auf. Es handelt sich somit nicht um einen separaten IR-Test, sondern um eine Nachfolgeuntersuchung der Relevanzbewertungen des vorherigen Tests. Basis dieser Evaluation sind die Dokumentenpools der jeweiligen Topics (siehe Abschnitte 5.4.1 und 5.4.2) und Testkollektionen (siehe Abschnitt 5.3).

Ziel der Untersuchung ist es, den Re-Ranking-Ansatz („Ranking nach Kernzeitschriften“ oder Bradfordizing) für Zeitschriftenartikel und Monographien anzuwenden und in einem realen Suchszenario zu evaluieren (siehe dazu Fragestellung 4 und 5 in Kapitel 1).

Die Hauptidee der Anwendung von Bradfordizing in dieser Arbeit ist es, ein beliebiges thematisch spezifiziertes Suchergebnis durch ein einfach anwendbares Verfahren umzuorganisieren (Re-Ranking) und den Effekt dieses Re-Rankings zu evaluieren (zur Verwendung des Begriffs Re-Ranking siehe (Bonitz, 1985); und aktuelles Experiment bei (Lin, 2008)). Der Einsatz von Bradfordizing soll das Suchergebnis komplett neu zusam-

mensetzen und möglichst das Suchergebnis signifikant verbessern. Ziel ist es, mehr relevante Dokumente für das Topic im ersten Drittel der Ergebnismenge zu listen. Die Anwendung von Bradfordizing bei der Suche stellt somit einen Mehrwert dar: Bradfordizing soll den Anteil der relevanten Dokumente im vorderen Trefferlistenbereich erhöhen und insgesamt den Suchprozess durch die neue Struktur des Suchergebnisses positiv beeinflussen. Das Listen der Kernzeitschriften vor den nachfolgenden und Peripherie-Zeitschriften eines Topics soll dem Nutzer die Möglichkeit geben, das Suchergebnis über diesen alternativen Zugangsmechanismus zu erkunden.

Die wichtigste Fragestellung des Re-Ranking-Experiments lautet:

Verbessert das Bradfordizing den Anteil der relevanten Dokumente im ersten Drittel (Core) einer Dokumentmenge gegenüber dem zweiten und dritten Drittel (Zone 2 bzw. Zone 3)?

Im Experiment wurden zunächst die beiden Dokumenttypen Zeitschriftenartikel und Monographien pro Topic aus den Dokumentenpools extrahiert und jeweils getrennt voneinander betrachtet (siehe Abbildung 5-5). Untersuchungsgegenstand dieser Evaluation sind die bewerteten Dokumentenpools (gepoolte Dateien) des IR-Tests im vorherigen Abschnitt. Aufgrund von zu geringen Dokumentzahlen konnten von den ursprünglichen 82 KoMoHe-Dokumentenpools insgesamt nur 39 Topics und Dokumente aus 6 unterschiedlichen Datenbanken in die Re-Ranking-Analyse einbezogen werden. Das Bradfordizing einer themenspezifischen Dokumentmenge sollte auf mindestens 100 Dokumenten (Zeitschriftenartikel oder Monographien) aufsetzen. Viele Topics der KoMoHe-Evaluation ergaben aber deutlich unter 100 Dokumente und wurden daher von der weiteren Analyse ausgenommen. Zusätzlich wurden 125 Topics des CLEF-Evaluationsprojekts analysiert. Eine Übersicht der Größe und Zusammensetzung der analysierten Dokumentenpools findet sich in Abschnitt 5.3.1.

Abbildung 5-5 verdeutlicht den Prozess der Auswertung der evaluierten Topics aus den beiden Evaluationsprojekten KoMoHe und CLEF nach dem Verfahren Bradfordizing.

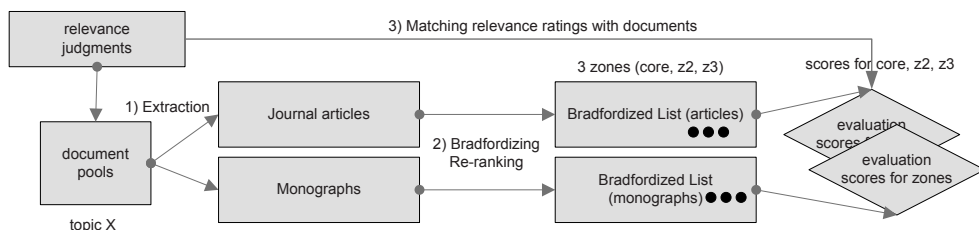


Abbildung 5-5: Auswertung der Relevanzbewertung im Re-Ranking-Experiment (KoMoHe- und CLEF-Topics)

Die Analyse der Datensätze teilt sich in drei Schritte:

1. Die Dokumenttypen Zeitschriftenartikel (journal articles) und Monographien (monographs) werden aus dem ursprünglichen Datensatz extrahiert. Jedes Topic wird separat analysiert.

- 2. Anschließend wird die Dokumentmenge pro Topic nach dem Verfahren Bradfordizing (Häufigkeit einer Zeitschrift bzw. eines Verlages) umorganisiert und in drei etwa gleich große Zonen eingeteilt (core, z2 und z3).
- 3. Im nächsten Schritt werden die Relevanzbewertungen der Dokumente in diesen Zonen analysiert (siehe Abbildung 5-6). Dazu werden die Dokumenten-IDs mit den Bewertungen zu den jeweiligen Topics abgeglichen und ausgewertet.

Im Anschluss an die Untersuchung liegen für jedes Topic und jeden Dokumenttyp die Anteile der relevanten bzw. nichtrelevanten Dokumente für die jeweilige Zone des Dokumentenpools vor. Daraus lassen sich Precision-Werte berechnen (siehe Abbildung 5-6).

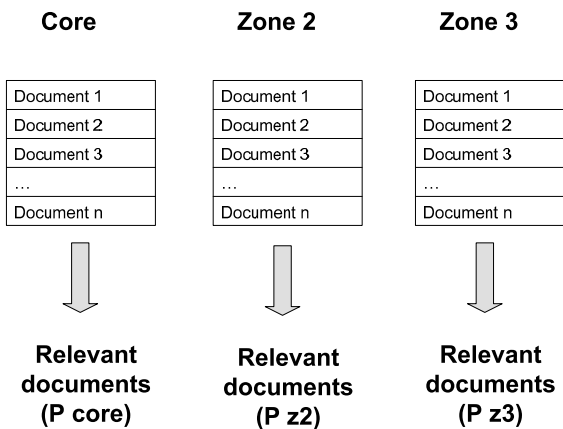


Abbildung 5-6: Aufteilung des Dokumentenpools pro Topic in drei gleich große Untermengen (Core, Zone 2 und Zone 3) und Abgleich der relevanten Dokumente pro Zone.

Wenn der Anteil der relevanten Dokumente, gemessen in Precision (P), in allen drei Zonen gleich groß ist, dann hat das Verfahren Bradfordizing keinen Effekt auf die Verteilung der Relevanzen im gesamten Dokumentpool. Wenn der Anteil der relevanten Dokumente in Zone Core niedriger ist als in Zone 2 oder Zone 3, dann verschlechtert sich das Suchergebnis. Wenn der Anteil der relevanten Dokumente im Core höher liegt als in den übrigen Zonen, was erwartet wird, dann verbessert das Verfahren das Suchergebnis und hat somit einen positiven Effekt auf die Suche.

Die Berechnung der Precision- und Recall-Werte wird in Abschnitt 5.7 erklärt.

5.3 Testkorpora

In diesem Kapitel soll die Zusammensetzung der einzelnen Testkorpora der beiden Evaluationsprojekte beschrieben werden.

5.3.1 GIRT-Korpus

GIRT (German Indexing and Retrieval Testdatabase) wird im Rahmen des domain-specific track der CLEF-Evaluation seit mehreren Jahren für Retrieval-Tests eingesetzt. Der GIRT-Korpus umfasst aktuell 151.319 deutschsprachige Metadaten zu Forschungsprojekten der Datenbank SOFIS und Literaturnachweisen der Datenbank SOLIS (Zeitraum 1990 – 2000). Für die Evaluation des Re-Rankings nach Bradfordizing wurden ausschließlich Zeitschriftenartikel und Monographien der Fachdatenbank SOLIS verwendet.

Das sozialwissenschaftliche Literaturinformationssystem (SOLIS⁶⁴) erfasst deutschsprachige fachwissenschaftliche Literatur (hauptsächlich Aufsätze in Zeitschriften, Beiträge in Sammelwerken, Monographien und Graue Literatur) aus allen Bereichen der Sozialwissenschaften. SOLIS erschließt mehr als 370.000 Veröffentlichungen (Stand Mai 2008). Die Nachweise in SOLIS werden mit dem Thesaurus Sozialwissenschaften erschlossen, der aktuell ca. 7.700 Deskriptoren umfasst.

Das sozialwissenschaftliche Forschungsinformationssystem (SOFIS⁶⁵) enthält ausführliche Beschreibungen von geplanten, laufenden und in den letzten zehn Jahren abgeschlossenen Forschungsarbeiten aus der Bundesrepublik Deutschland, aus Österreich und der Schweiz (Bestand im Frühsommer 2008: über 44.500 Dokumente). Pro Jahr werden 6.000 bis 7.000 Projektbeschreibungen neu oder in aktualisierter Version in die Datenbank aufgenommen. Die Nachweise in SOFIS werden ebenfalls mit dem Thesaurus Sozialwissenschaften erschlossen.

64 <http://www.gesis.org/dienstleistungen/fachinformationen/datenbankeninformationssysteme/literaturdatenbank-solis/>

65 <http://www.gesis.org/dienstleistungen/fachinformationen/datenbankeninformationssysteme/forschungsdatenbank-sofis/>

```

43 <DOC>
44 <DOCID>iz-solis-90123690</DOCID>
45 <IDENTIFIER1>19900100587</IDENTIFIER1>
46 <ISSN>0044-3360</ISSN>
47 <TITLE-DE>Politik ohne Vertrauen?</TITLE-DE>
48 <DOCTYPE>journalarticle</DOCTYPE>
49 <SOURCE>Zeitschrift fuer Politik Organ der Hochschule fuer Politik Muenchen, N. F.</SOURCE>
50 <MEDIATYPE/>
51 <AUTHOR>Backes, Uwe</AUTHOR>
52 <PUBLICATION-YEAR>1990</PUBLICATION-YEAR>
53 <LANGUAGE-CODE>de</LANGUAGE-CODE>
54 <CONTROLLED-TERM-DE>Bericht#Gesellschaft#Industriestaat#Inhalt#Konferenz#Osteuropa#
55 Politik#Politikwissenschaft#Vertrauen#westliche Welt#20. Jahrhundert</CONTROLLED-TERM-DE>
56 <ABSTRACT-DE>Der Beitrag berichtet ueber die Jahrestagung der Deutschen Gesellschaft fuer
57 Politikwissenschaft am 17. und 18.11.1989 in Erfstadt bei Koeln und fasst die wesentlichen
58 Inhalte der Referate zusammen. Leitthema war die Frage 'Politik ohne Vertrauen?', die sich zum
59 einen auf den diagnostizierten Vertrauensverlust demokratischer Institutionen und politischer
60 Parteien bezog, zum anderen durch die Wandlungsprozesse in den Staaten Osteuropas inspiriert
61 war. Die Tagung wurde eroeffnet mit Referaten zur 'grundsaeztlichen Bedeutung der Kategorien
62 'Vertrauen' und 'Misstrauen' fuer das Funktionieren politischer Systeme im allgemeinen,
63 konstitutioneller Demokratien im besonderen'. Es folgten Ausfuehrungen zum messbaren Ausmass
64 an Vertrauen bzw. Misstrauen der Bevoelkerung in westlichen Laendern gegenueber ihren
65 politischen Repraesentanten, zur Rolle des Vertrauens in der internationalen Politik, zu
66 Vertrauensbildung, Konfliktschaerfung und Friedensstiftung bei den Voelkern des Nahen
67 Ostens. Abschliessend wurden am Beispiel Polens drei Dimensionen des Vertrauens in den
68 Laendern Osteuropas untersucht: gemeinsame Werte als verbindlich geltende Ideologie,
69 politische Prozesse der Parteiherrschaft und Effizienz der Parteiherrschaft. Auf diese drei
70 Dimensionen erstreckte sich Vertrauen bzw. Misstrauen. (AG)</ABSTRACT-DE>
71 <TEXT/>
72 </DOC>

```

Abbildung 5-7: Typisches Dokument im erweiterten GIRT4-Format

Abbildung 5-7 zeigt die Metadaten eines Zeitschriftenartikels im erweiterten GIRT4-Format. Zusätzliche Elemente gegenüber dem GIRT4-Format⁶⁶ sind die Felder DOCTYPE, ISSN und SOURCE. Da diese Elemente (insbesondere DOCTYPE und ISSN) notwendig für die spätere Häufigkeitsanalyse sind, wurden sie vor der Bradfordizing-Analyse zu den Dokumenten hinzugefügt.

5.3.1 KoMoHe-Kollektionen

Die Dokumente der KoMoHe-Evaluation wurden ebenfalls alle in das oben dargestellte erweiterte GIRT4-Format konvertiert. Die in den Experimenten genutzten bibliographischen Datenbanken werden hauptsächlich in Deutschland produziert und gehostet und enthalten zwischen ca. 70.000 - 16 Millionen Dokumente. Unter den Datenbanken befinden sich Fachdatenbanken, aber auch Bibliothekskataloge. Tabelle 15 gibt einen Überblick über die getesteten Datenbanken und ihre assoziierten kontrollierten Vokabulare:

66 <http://www.gesis.org/forschung-lehre/programme-projekte/informationswissenschaften/projektuebersicht/clef/cross-language-domain-specific-information-retrieval-girt4/>

Tabelle 15: Vokabulare und Datenbanken im KoMoHe IR-Test

| Vokabular | Disziplin | Datenbank | Dokumente in DB (Stand: August 2007) ⁷² |
|---|---|--|---|
| TheSoz – Thesaurus Sozialwissenschaften (GESIS) | Sozialwissenschaften | SOLIS (in-house) | 345.086 |
| DZI – Thesaurus des Deutschen Instituts für soziale Fragen | Sozialwissenschaften | SoLit (in-house) | 151.925 |
| SWD – Schlagwortnormdatei | übergreifend (sozialwissenschaftlicher Ausschnitt) | USB Köln Sowi OPAC (in-house) | 72.729 |
| CSA – Thesaurus of Sociological Indexing Terms (Cambridge Scientific Abstracts) | Sozialwissenschaften | CSA Sociological Abstracts ⁷³ (in-house) | 294.875 |
| Psyndex – Psyndex Terms | Psychologie | Psyndex (ZPID) | ca. 200.000 |
| STW – Standard Thesaurus Wirtschaft | Wirtschaftswissenschaften | Econis (ZBW Kiel) | ca. 3.000.000 |
| IBLK – Thesaurus Internationale Beziehungen und Länderkunde (Euro-Thesaurus) | Politikwissenschaft | World Affairs Online WAO (SWP Berlin) (in-house) | 643.420 |
| Mesh – Medical Subject Headings | Medizin | Medline (Dimdi) | ca. 16.800.000 |

Im Folgenden werden die analysierten Topics der beiden Evaluationsprojekte CLEF und KoMoHe beschrieben.

5.4 Topics

Für diese Studie wurden Topics aus zwei Evaluationsprojekten analysiert. Tabelle 16 liefert eine Übersicht der bewerteten Dokumente aus beiden Datensätzen. Insgesamt wurden im Rahmen der 125 CLEF-Topics 65.297 Dokumente und aus dem KoMoHe-Datensatz 31.155 Dokumente bewertet. In die Bradfordizing-Analyse konnten 29.157 (CLEF) und 22.332 (KoMoHe) Dokumente einbezogen werden. Für die Bradfordizing-Analyse konnten aus dem CLEF-Datensatz nur die Literaturnachweise aus der Datenbank SOLIS berücksichtigt werden. Die Nachweise von SOFIS (Forschungsprojekte) wurden bei der Analyse ausgeklammert.

67 Im August wurden die Dokumente für den IR-Test in das Testsystem gespielt.

68 Die Dokumente der Datenbank CSA-SA konnten aus Zeitgründen nicht mehr bewertet werden.

Tabelle 16: Übersicht über die analysierten Topics, Dokumente und Datenbanken

| | CLEF | KoMoHe |
|-----------------------------------|-----------|--------|
| Zeitraum | 2003-2007 | 2007 |
| Anzahl der Topics | 125 | 39 |
| bewertete Dokumente total | 65.297 | 31.155 |
| Dokumente bradfordized | 29.157 | 22.332 |
| Zeitschriftenartikel bradfordized | 18.112 | 17.432 |
| Monographien bradfordized | 11.045 | 4.900 |
| Datenbanken involviert | 2 (1) | 6 |

5.4.1 CLEF-Topics

Die einzelnen Topics sind standardisiert aufgebaut und enthalten eine möglichst klar umrissene Fragestellung. Jedes Topic erhält eine eindeutige Nummer, einen Titel, eine Kurzbeschreibung und eine längere Beschreibung (siehe Abbildung 5-8). Die Topics werden nach der Fertigstellung in die Sprachen der CLEF-Sprachgruppen übersetzt (Deutsch, Englisch, Russisch) und auf unterschiedliche Test-Kollektionen angesetzt. Bei den folgenden Topics handelt es sich um die deutschsprachigen Versionen, die in der GIRT-Kollektion abgefragt und analysiert wurden. Die Topics werden vor den Retrieval-Tests von den Organisatoren der Retrieval-Konferenz getestet und bearbeitet. Geprüft wird insbesondere, ob die Fragestellungen den Scope der Datenbank SOLIS abdecken. Es werden nur Topics berücksichtigt, die ausreichend Dokumente in der entsprechenden Kollektion generieren. Bei den folgenden Topics der Jahre 2003 - 2007 handelt es sich daher im weitesten Sinne um wissenschaftliche Fragestellungen in den Sozialwissenschaften.

```
<top><num> 091 </num>
<DE-title> Qualität von Bildungssystemen </DE-title>
<DE-desc> Finde Dokumente, die die Sicherung der Qualität von
Bildungssystemen in verschiedenen Staaten thematisieren. </DE-desc>
<DE-narr> Relevante Dokumente berichten über Modelle, politische Ziele
oder theoretische Ansätze zur Sicherung der Qualität des Outputs von
Bildungssystemen. Hierbei sind zum Beispiel Maßnahmen zum prozessualen
Controlling oder Finanzcontrolling, Evaluationsmodelle oder Systeme der
Netzwerkbildung zur Kooperation von an Bildung beteiligten Institutionen
von Interesse. Wichtig ist, dass die Dokumente qualitative Aspekte der
Bildung behandeln. </DE-narr>
</top>
```

Abbildung 5-8: Typisches Topic mit den Elementen: Nummer (num), dt. Titel (DE-title), dt. Kurzbeschreibung (DE-desc) und dt. Langbeschreibung (DE-narr).

“Topics for retrieval are written by people who are actually users of a retrieval system and devised to mimic real users’ needs. The topics are constructed by doing trial retrievals against a sample of the document set, and then those topics that have roughly 25 to 100 hits in that sample are selected.” ((Chu, 2003) S. 214)

Die folgenden Übersichtstabellen zeigen die Titel und Nummern der analysierten Topics der CLEF-Jahrgänge 2003 - 2007 sowie die jeweiligen Dokumentzahlen im GIRT-Korpus. Aus zeitlichen Gründen wurden für die CLEF-Topics der Jahre 2003 und 2004 lediglich Zeitschriftenartikel analysiert. Für die Jahrgänge 2005, 2006 und 2007 wurden zusätzlich die Monographien in die Analyse einbezogen.

Tabelle 17: Topic-Titel, Topic-Nr., Gesamtzahl der gepoolten Dokumente (Dok. total), Anzahl der Zeitschriftenartikel (Zt.art.) und restliche Dokumente (Rest) pro Topic des domain-specific CLEF track 2003

| Titel | Nr. | Dok. total | Zt.art. | Rest |
|--|-----|------------|---------|------|
| Politische Kultur in einer Demokratie | 76 | 459 | 121 | 338 |
| Politische Partizipation von DDR-Frauen | 77 | 424 | 91 | 333 |
| Bildung in der Türkei | 78 | 404 | 128 | 276 |
| Volksentscheide in der Demokratie | 79 | 241 | 92 | 149 |
| Industrielle und ökonomische Entwicklung Indiens | 80 | 495 | 129 | 366 |
| Ausbildungsabbruch | 81 | 525 | 199 | 326 |
| Berufliche Bildung von Immigranten | 82 | 586 | 198 | 388 |
| Medien und Krieg | 83 | 406 | 187 | 219 |
| Neue Medien im Unterricht | 84 | 485 | 203 | 282 |
| Europäische Bildungssysteme im Vergleich | 85 | 442 | 95 | 347 |
| Politische Symbole in Osteuropa | 86 | 486 | 116 | 370 |
| Integration Behinderter in Schulen | 87 | 347 | 119 | 228 |
| Sport im Nationalsozialismus | 88 | 526 | 114 | 412 |
| Berufsethik im Journalismus | 89 | 416 | 170 | 246 |
| Zeitunglesen | 90 | 460 | 240 | 220 |
| Qualität von Bildungssystemen | 91 | 470 | 144 | 326 |
| Parteimitglieder | 92 | 412 | 133 | 279 |
| Burnout-Syndrom | 93 | 308 | 119 | 189 |
| Homosexualität und Coming-out | 94 | 353 | 120 | 233 |
| Sonntagsarbeit | 95 | 440 | 176 | 264 |
| Kosten der Berufsausbildung | 96 | 366 | 132 | 234 |
| Sekten | 97 | 494 | 131 | 363 |
| Außenpolitik Kanadas | 98 | 469 | 168 | 301 |
| Sport und Jugendarbeit | 99 | 459 | 102 | 357 |
| Kneipenkultur | 100 | 664 | 174 | 490 |

Tabelle 17 präsentiert die 25 Topics aus dem Jahr 2003. Im Rahmen der Evaluation wurden zu diesen Topics insgesamt 11.137 sozialwissenschaftliche Dokumente im GIRT-Korpus auf Relevanz bewertet. 3.601 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen. Im Mittel wurden 144 Zeitschriftenartikel pro Topic analysiert.

Tabelle 18: Topic-Titel, Topic-Nr., Gesamtzahl der gepoolten Dokumente (Dok. total), Anzahl der Zeitschriftenartikel (Zt.art.) und restliche Dokumente (Rest) pro Topic des domain-specific CLEF track 2004

| Titel | Nr. | Dok. total | Zt.art. | Rest |
|--|-----|------------|---------|------|
| Tiere in der Therapie | 101 | 378 | 118 | 260 |
| Deregulierung des Strommarktes | 102 | 279 | 84 | 195 |
| Wohnkultur | 103 | 485 | 65 | 420 |
| Prävention arbeitsbedingter Krankheiten | 104 | 370 | 98 | 272 |
| Akademiker und Arbeitsmarkt | 105 | 361 | 137 | 224 |
| Multiple Sklerose | 106 | 388 | 86 | 302 |
| Konfliktmanagement | 107 | 558 | 171 | 387 |
| Transportlogistik | 108 | 293 | 58 | 235 |
| Ökologische Landwirtschaft | 109 | 509 | 133 | 376 |
| Selbstmord von Jugendlichen | 110 | 315 | 92 | 223 |
| Elternarbeit in der Schule | 111 | 356 | 103 | 253 |
| Supermärkte | 112 | 357 | 161 | 196 |
| Katholizismus in Polen | 113 | 351 | 91 | 260 |
| Illegale Beschäftigung in der Bundesrepublik Deutschland | 114 | 535 | 204 | 331 |
| Ökonomische Auswirkungen der Ökosteuer | 115 | 306 | 118 | 188 |
| Prüfungsangst | 116 | 543 | 159 | 384 |
| Handel von Emissionsrechten | 117 | 330 | 86 | 244 |
| Religion und Politik | 118 | 378 | 79 | 299 |
| Assessment Center zur Personalrekrutierung | 119 | 302 | 127 | 175 |
| Städtebau der Nachkriegszeit | 120 | 395 | 115 | 280 |
| Corporate identity | 121 | 403 | 107 | 296 |
| Bevölkerungsstatistik für Deutschland | 122 | 474 | 154 | 320 |
| Körpertraining | 123 | 562 | 147 | 415 |
| Telefonumfragen | 124 | 246 | 79 | 167 |
| Säuglingssterblichkeit | 125 | 262 | 96 | 166 |

Tabelle 18 präsentiert die 25 Topics aus dem Jahr 2004. Im Rahmen der Evaluation wurden zu diesen Topics insgesamt 9.736 sozialwissenschaftliche Dokumente im GIRT-Korpus auf Relevanz bewertet. 2.868 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen. Im Mittel wurden 114 Zeitschriftenartikel pro Topic analysiert.

Tabelle 19: Topic-Titel, Topic-Nr., Gesamtzahl der gepoolten Dokumente (Dok. total), Anzahl der Zeitschriftenartikel (Zt.art.), Anzahl der Monographien (Monogr.) und restliche Dokumente (Rest) pro Topic des domain-specific CLEF track 2005

| Titel | Nr. | Dok. total | Zt.art. | Monogr. | Rest |
|---|-----|------------|---------|---------|------|
| Neue Kunst | 126 | 501 | 136 | 148 | 217 |
| Wahlverhalten | 127 | 612 | 170 | 146 | 296 |
| Lebenszufriedenheit | 128 | 556 | 118 | 145 | 293 |
| Sexualität und Behinderung | 129 | 695 | 207 | 218 | 270 |
| Wassermangel | 130 | 507 | 95 | 188 | 224 |
| Zweisprachige Erziehung | 131 | 630 | 157 | 166 | 307 |
| Sexueller Missbrauch von Kindern | 132 | 190 | 57 | 64 | 69 |
| Kirchen und Geld | 133 | 496 | 142 | 162 | 192 |
| Russisch-chinesische Wirtschaftsbeziehungen | 134 | 434 | 79 | 189 | 166 |
| Rente in post-sowjetischen Ländern | 135 | 663 | 190 | 203 | 270 |
| Ökologische Abfallwirtschaft | 136 | 628 | 104 | 240 | 284 |
| Ehre und Gesellschaft | 137 | 414 | 93 | 130 | 191 |
| Unternehmensinsolvenzen | 138 | 904 | 208 | 264 | 432 |
| Gesundheitsökonomie | 139 | 603 | 94 | 180 | 329 |
| Erdöl und Politik | 140 | 675 | 147 | 279 | 249 |
| Straßenkinder | 141 | 499 | 101 | 163 | 235 |
| Werbung und Ethik | 142 | 535 | 240 | 92 | 203 |
| Aufgeben des Rauchens | 143 | 296 | 100 | 63 | 133 |
| Radio und Internet | 144 | 430 | 247 | 43 | 140 |
| Armut und Reichtum | 145 | 352 | 105 | 85 | 162 |
| Diabetes Mellitus | 146 | 428 | 92 | 111 | 225 |
| Fußball und Gesellschaft | 147 | 549 | 160 | 124 | 265 |
| Russlanddeutsche und Sprache | 148 | 508 | 103 | 188 | 217 |
| Antisemitismus in der Sowjetunion | 149 | 515 | 110 | 194 | 211 |
| Fernsehverhalten | 150 | 568 | 301 | 61 | 206 |

Tabelle 19 präsentiert die 25 Topics aus dem Jahr 2005. Im Rahmen der Evaluation wurden zu diesen Topics insgesamt 13.188 sozialwissenschaftliche Dokumente im GIRT-Korpus auf Relevanz bewertet. 3.556 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen. Zudem wurden 3.846 Monographien ausgewertet. Im Mittel wurden 142 Zeitschriftenartikel und 153 Monographien pro Topic analysiert.

Tabelle 20: Topic-Titel, Topic-Nr., Gesamtzahl der gepoolten Dokumente (Dok. total), Anzahl der Zeitschriftenartikel (Zt.art.), Anzahl der Monographien (Monogr.) und restliche Dokumente (Rest) pro Topic des domain-specific CLEF track 2006

| Titel | Nr. | Dok. total | Zt.art. | Monogr. | Rest |
|---|-----|------------|---------|---------|------|
| Rechtsextreme Parteien in Deutschland | 151 | 533 | 128 | 151 | 254 |
| Beschäftigungspolitik auf europäischer Ebene | 152 | 917 | 292 | 281 | 344 |
| Kinderlosigkeit in Deutschland | 153 | 428 | 133 | 70 | 225 |
| Modernisierung der öffentlichen Verwaltung | 154 | 500 | 75 | 140 | 285 |
| Häusliche Gewalt | 155 | 681 | 166 | 160 | 355 |
| Illegale Aufenthalte in Deutschland | 156 | 854 | 288 | 177 | 389 |
| Mehrsprachigkeit von Kindern | 157 | 495 | 143 | 86 | 266 |
| Rückwanderung und Transmigration | 158 | 712 | 163 | 178 | 371 |
| Vaterrolle | 159 | 361 | 105 | 66 | 190 |
| Prekäre Arbeitsverhältnisse | 160 | 782 | 205 | 118 | 459 |
| Europäische Sozialpolitik | 161 | 613 | 156 | 204 | 253 |
| Mutterschaft und berufliche Entwicklung | 162 | 687 | 146 | 147 | 394 |
| Risikoverhalten von Kindern und Jugendlichen | 163 | 738 | 197 | 130 | 411 |
| Das deutsche Schulsystem | 164 | 573 | 162 | 123 | 288 |
| Kinderbetreuung durch Institutionen | 165 | 661 | 200 | 147 | 314 |
| Armut in Deutschland | 166 | 532 | 125 | 127 | 280 |
| Kriminalität bei Frauen | 167 | 766 | 159 | 179 | 428 |
| Antisemitismus in Deutschland nach 1945 | 168 | 486 | 64 | 197 | 225 |
| Geschlechtsspezifische Lernformen in der Grundschule | 169 | 620 | 152 | 100 | 368 |
| Lean-production in Japan | 170 | 466 | 117 | 148 | 201 |
| Computer im Alltag | 171 | 585 | 208 | 113 | 264 |
| Ausländer in der Grundschule | 172 | 440 | 109 | 64 | 267 |
| Gewaltbereitschaft von Jugendlichen | 173 | 470 | 99 | 61 | 310 |
| Befragung von Kindern | 174 | 429 | 89 | 122 | 218 |
| Bildung der Eltern und die schulische Entwicklung von Kindern | 175 | 619 | 148 | 99 | 372 |

Tabelle 20 präsentiert die 25 Topics aus dem Jahr 2006. Im Rahmen der Evaluation wurden zu diesen Topics insgesamt 14.948 sozialwissenschaftliche Dokumente im GIRT-Korpus auf Relevanz bewertet. 3.829 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen. Zudem wurden 3.388 Monographien ausgewertet. Im Mittel wurden 153 Zeitschriftenartikel und 135 Monographien pro Topic analysiert.

Tabelle 21: Topic-Titel, Topic-Nr., Gesamtzahl der gepoolten Dokumente (Dok. total), Anzahl der Zeitschriftenartikel (Zt.art.), Anzahl der Monographien (Monogr.) und restliche Dokumente (Rest) pro Topic des domain-specific CLEF track 2007

| Titel | Nr. | Dok. total | Zt.art. | Monogr. | Rest |
|--|-----|------------|---------|---------|------|
| Geschwisterbeziehungen | 176 | 703 | 161 | 163 | 379 |
| Arbeitslose Jugendliche ohne Berufsausbildung | 177 | 547 | 219 | 85 | 243 |
| Deutsch-französische Beziehungen nach 1945 | 178 | 596 | 125 | 325 | 146 |
| Multinationale Unternehmen | 179 | 392 | 102 | 116 | 174 |
| Partnerschaft und Kinderwunsch | 180 | 584 | 173 | 72 | 339 |
| Folter im Rechtsstaat | 181 | 573 | 141 | 249 | 183 |
| Familienpolitik und Volkswirtschaft | 182 | 659 | 129 | 120 | 410 |
| Frauen und Gehaltsniveau | 183 | 783 | 383 | 94 | 306 |
| Lebensstil und Umweltverhalten | 184 | 456 | 65 | 89 | 302 |
| Unsichere Erwerbsarbeitsverhältnisse | 185 | 1161 | 319 | 186 | 656 |
| Wertewandel in Osteuropa | 186 | 855 | 203 | 217 | 435 |
| Migrationsdruck | 187 | 578 | 187 | 108 | 283 |
| Lebensqualität älterer Menschen | 188 | 538 | 99 | 108 | 331 |
| Schichtspezifisches Freizeitverhalten | 189 | 540 | 71 | 51 | 418 |
| Sterberate | 190 | 565 | 249 | 78 | 238 |
| Wirtschaftseliten in Osteuropa und Rußland | 191 | 538 | 98 | 181 | 259 |
| Systemwandel und Familienbildung in Ostdeutschland | 192 | 696 | 171 | 106 | 419 |
| Geschlecht und Karrierechancen | 193 | 829 | 208 | 154 | 467 |
| Ökostandards in Schwellenländern | 194 | 776 | 197 | 212 | 367 |
| Integrationspolitik | 195 | 774 | 199 | 178 | 397 |
| Tourismusbranche in Deutschland | 196 | 622 | 145 | 192 | 285 |
| Betriebliche Gesundheitsförderung | 197 | 420 | 141 | 63 | 216 |
| Ökonomische Situation von Familien | 198 | 578 | 120 | 92 | 366 |
| Europäische Klimapolitik | 199 | 581 | 85 | 311 | 185 |
| Wirtschaftsförderung im Osten | 200 | 944 | 268 | 261 | 415 |

Tabelle 21 präsentiert die 25 Topics aus dem Jahr 2007. Im Rahmen der Evaluation wurden zu diesen Topics insgesamt 16.288 sozialwissenschaftliche Dokumente im GIRT-Korpus auf Relevanz bewertet. 4.258 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen. Zudem wurden 3.811 Monographien ausgewertet. Im Mittel wurden 170 Zeitschriftenartikel und 152 Monographien pro Topic analysiert.

5.4.2 KoMoHe-Topics

Bei den folgenden Topics im Projekt KoMoHe handelt es sich um typische Rechercheanfragen, die bei den Partnerinstitutionen der GESIS in der Vergangenheit eingegangen sind und von den jeweiligen Informationsspezialisten bearbeitet wurden. Die Partnerinstitutionen wurden gebeten, den ursprünglichen natürlichsprachigen Rechercheantrag sowie die Operationalisierung (möglichst in Deskriptoren) der Fachrecherchen zur Verfügung zu stellen. Aus der Vielzahl der themenspezifischen Fragestellungen wurde ein Set von Anfragen ausgewählt und für den Retrieval-Test operationalisiert. Zum Teil konnten die ursprünglichen Recherchestrategien kaum verändert übernommen werden, weil die Anfragen ausschließlich kontrollierte Terme enthielten. Die Topics im Projekt KoMoHe wurden vor den Retrieval-Tests wie in den klassischen Retrieval-Konferenzen von den Projektmitarbeitern getestet und ggf. bearbeitet. Geprüft wurde insbesondere, ob die Fragestellungen den Scope der Datenbanken abdecken. Es werden nur Topics berücksichtigt, die ausreichend Dokumente in den entsprechenden Kollektionen generiert haben.

Die 39 Topics lassen sich drei unterschiedlichen Tests zuordnen:

- Test 1 beinhaltet 12 Topics, die Fragestellungen in den Sozialwissenschaften behandeln. Es werden Dokumente der sozialwissenschaftlichen Datenbanken (SOLIS, SOLIT und USB-Köln OPAC) untersucht. Dieser Test wird als innerdisziplinärer Test (Test 1) bezeichnet.
- Test 2 beinhaltet 15 Topics, die Fragestellungen in den Sozialwissenschaften und verwandten Disziplinen behandeln. Es werden Dokumente der Datenbanken (SOLIS, Psynindex und WAO) untersucht. Dieser Test wird als interdisziplinärer Test (Test 2) bezeichnet.
- Test 3 beinhaltet 12 Topics, die Fragestellungen in Medizin und Psychologie (außerhalb der Sozialwissenschaften) behandeln. Es werden Dokumente der Datenbanken (Psynindex und Medline) untersucht. Dieser Test wird als interdisziplinärer Test (Test 3) bezeichnet.

Die folgenden Übersichtstabellen zeigen die Titel und Nummern der analysierten Topics und die jeweiligen Dokumentzahlen in den Zieldatenbanken der KoMoHe-Evaluation. Für die KoMoHe-Topics wurden wenn möglich Zeitschriftenartikel und Monographien analysiert. Einige Topics enthielten keine Monographien, dies wird in den Tabellen angegeben. Tabelle 22 und Tabelle 23 präsentieren die 39 Topics der KoMoHe-Evaluation 2007, die in die Bradford-Analyse einbezogen werden konnten. Einige Topics tragen

identische Titel. Da diese Topics im IR-Test in unterschiedlichen Datenbanken evaluiert werden, kann diese Tatsache aber vernachlässigt werden.

Tabelle 22: Topic-Titel, Topic-Nr., Testreihe, Zieldatenbank (Ziel-DB), Gesamtzahl der gepoolten Dokumente (Dok. total) und Anzahl der Zeitschriftenartikel (Zt.art.) und Monographien (Monogr.) und die restlichen Dokumente (Rest) pro Topic der KoMoHe-Evaluation 2007 (Datenbanken: SOLIS, SOLIT, USB Köln und WAO)

| Titel | Nr. | Test | Ziel-DB | Dok. total | Zt.art. | Monogr. | Rest |
|--|------|--------|----------|------------|---------|---------|------|
| Family life (Familienleben) | 513 | Test 1 | SOLIS | 1422 | 488 | 315 | 619 |
| Deviant Behavior in adolescent age (Abweichendes Verhalten von Jugendlichen) | 515 | Test 1 | SOLIS | 519 | 139 | 124 | 256 |
| Ethnicity (Ethnizität) | 518 | Test 1 | SOLIS | 1380 | 289 | 477 | 614 |
| Bedeutung der Herkunftsfamilie für den Schulerfolg | 414 | Test 1 | SOLIS | 218 | 60 | 49 | 109 |
| Auswirkungen von Adoption/ Dauerpflege | 311 | Test 2 | SOLIS | 215 | 87 | 46 | 82 |
| Effektivität/Wirksamkeit von Psychotherapien im Allgemeinen | 313 | Test 2 | SOLIS | 933 | 365 | 236 | 332 |
| Behandlung von Drogenabhängigen und Traumatisierung | 314 | Test 2 | SOLIS | 821 | 257 | 230 | 334 |
| Stand der Forschung zur Erweiterung des Arbeitsbegriffes | 1312 | Test 1 | SOLIS | 580 | 148 | 167 | 265 |
| Stellung der Frau in der Sozialversicherung | 1313 | Test 1 | SOLIS | 607 | 253 | 103 | 251 |
| Auswirkungen des demographischen Wandels auf das Ausbildungswesen | 1314 | Test 1 | SOLIS | 467 | 172 | 70 | 225 |
| Systemtheorie | 140 | Test 1 | SOLIT | 585 | 437 | | 148 |
| Gewalt gegen Frauen | 141 | Test 1 | SOLIT | 667 | 430 | | 237 |
| Gesundheitsreform | 144 | Test 1 | SOLIT | 1080 | 883 | | 197 |
| Wertewandel | 148 | Test 1 | SOLIT | 1132 | 964 | | 168 |
| Arbeitsvermittlung | 149 | Test 1 | SOLIT | 1517 | 988 | | 529 |
| Systemtheorie | 1130 | Test 2 | USB-Köln | 1003 | | 665 | 338 |
| Unterschicht in Deutschland | 1132 | Test 2 | USB-Köln | 389 | | 266 | 123 |
| Wertewandel | 1138 | Test 2 | USB-Köln | 435 | | 299 | 136 |
| Mitbestimmung in Deutschland | 8910 | Test 2 | WAO | 470 | 164 | 116 | 190 |
| Informationen über das TRIPS-Abkommen | 8911 | Test 2 | WAO | 498 | 236 | 90 | 172 |

| Titel | Nr. | Test | Ziel-DB | Dok. total | Zt.art. | Monogr. | Rest |
|--|-----|--------|---------|------------|---------|---------|------|
| Aktuelle Entwicklungstendenzen im E-Business | 895 | Test 2 | WAO | 424 | 157 | 96 | 171 |
| Kostenrechnung | 898 | Test 2 | WAO | 351 | 120 | 63 | 168 |
| NAFTA (North American Free Trade Agreement) | 899 | Test 2 | WAO | 1341 | 674 | 247 | 420 |

Im Rahmen der Evaluation wurden zu 23 Topics insgesamt 17.054 Dokumente in den Datenbanken (SOLIS, SOLIT, USB-Köln und WAO) auf Relevanz bewertet. 7.311 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen. Zudem wurden 3.659 Monographien ausgewertet. Im Mittel wurden 365 Zeitschriftenartikel und 203 Monographien pro Topic analysiert.

Tabelle 23: Topic-Titel, Topic-Nr., Testreihe, Zieldatenbank (Ziel-DB), Gesamtzahl der gepoolten Dokumente (Dok. total) und Anzahl der Zeitschriftenartikel (Zt.art.) und Monographien (Monogr.) und die restlichen Dokumente (Rest) pro Topic der KoMoHe-Evaluation 2007 (Datenbanken: Medline und Psyn dex)

| Titel | Nr. | Test | Ziel DB | Dok. total | Zt.art. | Monogr. | Rest |
|---|-----|--------|----------|------------|---------|---------|------|
| Dementieller Sprachabbau bei Alzheimer-Demenz | 361 | Test 3 | MEDLINE | 836 | 836 | | |
| Ehescheidungen bzw. Ehestabilität | 362 | Test 3 | MEDLINE | 1041 | 1041 | | |
| Bewegungstherapie | 363 | Test 3 | MEDLINE | 1995 | 1995 | | |
| Ambulante Psychotherapie | 364 | Test 3 | MEDLINE | 636 | 636 | | |
| Lebensstil und Lebensqualität | 365 | Test 3 | MEDLINE | 1088 | 1088 | | |
| Paar- und Partnertherapie | 366 | Test 3 | MEDLINE | 1145 | 1145 | | |
| Schizophrenie | 367 | Test 3 | MEDLINE | 380 | 380 | | |
| Ehescheidungen bzw. Ehestabilität | 632 | Test 3 | PSYINDEX | 411 | 154 | 112 | 145 |
| Bewegungstherapie | 633 | Test 3 | PSYINDEX | 535 | 267 | 139 | 129 |
| Ambulante Psychotherapie | 634 | Test 3 | PSYINDEX | 396 | 214 | 53 | 129 |
| Paar- und Partnertherapie | 636 | Test 3 | PSYINDEX | 766 | 301 | 180 | 285 |
| Schizophrenie | 637 | Test 3 | PSYINDEX | 538 | 204 | 99 | 235 |
| Gewalt gegen Frauen | 131 | Test 2 | PSYINDEX | 779 | 317 | 149 | 313 |
| Unterschicht in Deutschland | 132 | Test 2 | PSYINDEX | 644 | 377 | 74 | 193 |
| Gewalt in der Schule | 133 | Test 2 | PSYINDEX | 703 | 261 | 188 | 254 |
| Gesundheitsreform | 134 | Test 2 | PSYINDEX | 1622 | 905 | 247 | 470 |

Im Rahmen der Evaluation wurden zu diesen 16 Topics insgesamt 13.514 Dokumente in den Datenbanken (Medline und Psyn dex) auf Relevanz bewertet. 10.121 Dokumente sind als Zeitschriftenartikel ausgezeichnet und wurden in die weitere Analyse einbezogen.

Zudem wurden 1.241 Monographien ausgewertet. Im Mittel wurden 632 Zeitschriftenartikel und 137 Monographien pro Topic analysiert.

5.5 Pooling der Dokumente

Wie bereits in den vorherigen Abschnitten erwähnt, basiert die Analyse der Dokumente auf gepoolten Datensätzen der beiden Evaluationsprojekte (CLEF und KoMoHe).

Pooling ist eine Methode, um den intellektuellen Bewertungsaufwand für eine Fragestellung zu minimieren. Ergebnismengen unterschiedlicher Systeme oder Systemläufe werden dabei nicht separat für sich evaluiert, sondern vor der Bewertung in eine gemeinsame Ergebnisliste pro Topic integriert (pooling). Die Juroren bewerten dann in einem Bewertungsdurchgang lediglich ein gepooltes (gemeinsames) Ergebnis, das sich aus den Einzelergebnissen aller Systeme im Test zusammensetzt. Bei den klassischen IR-Tests wie TREC oder CLEF liefern die Teilnehmer der Evaluierungsrunde jeweils ihre nach Relevanz gerankten Ergebnislisten pro Topic an den Evaluationsorganisator. Aus diesen Ergebnislisten werden die top-gerankten Dokumente (meist 60 bis 100 Treffer) pro Teilnehmer und Topic entnommen (siehe Abbildung 5-2). Aus diesen Dokumenten wird im nächsten Schritt ein neuer gemeinsamer Dokumentenpool zusammengestellt und in eine zufällige Reihenfolge gebracht. Zum Schluss existiert ein dublettenfreier Pool an Dokumenten zu einem Topic, der von den Juroren bezüglich Relevanz bewertet werden kann.

In dieser Studie wurde z.B. auf die gepoolten Dokumente der fünf Jahrgänge des domain-specific track der CLEF-Initiative der Jahre 2003 - 2007 zurückgegriffen. Eine Übersicht über die Anzahl der gepoolten Dokumente befindet sich in den Tabellen (Tabelle 17 - Tabelle 21). Die Dokumentenpools dieser fünf Jahrgänge umfassten im Mittel 522 Dokumente pro Topic⁶⁹.

Für das Pooling der Dokumente im Projekt KoMoHe wurden die drei Anfragetypen (CT, TT und FT, vgl. Abschnitt 5.2.1) und die resultierenden Suchergebnisse dieser drei unterschiedlichen Anfragen pro Topic verwendet. Für jeden Anfragetyp wurden maximal die ersten 1.000 Dokumente in den Dokumentenpool gespielt und anschließend gepoolt. Die Dokumentenpools der KoMoHe-Topics umfassten im Mittel 783 Dokumente pro Topic. Damit liegen die Dokumentenzahlen der KoMoHe-Evaluation um etwa 50 % über den durchschnittlichen Dokumentenzahlen der CLEF-Evaluation.

5.6 Relevanz-Bewertung

Die gepoolten Dokumente pro Topic werden den Juroren über ein Tool zur Bewertung bereitgestellt (vgl. Abbildung 5-9). **Daraufhin werden die Dokumente einzeln bezüglich der Relevanz⁷⁰ zu dem entsprechenden Topic bewertet** (siehe dazu (Harter, 1996)). Die Bewertung der Dokumente durch die Juroren erfolgt binär: Ein Dokument kann

69 In diesen Zahlen sind die Forschungsprojektnachweise aus SOFIS mit subsumiert, die später für die Analyse nicht berücksichtigt wurden.

70 Siehe dazu Definition bei Saracevic: "Relevance is the correspondence in context between an information requirement statement and an article; i.e., the extent to which the article covers

aufgrund seiner Merkmale als relevant (1) oder als nichtrelevant (0) zu einem Topic bewertet werden. Abgestufte oder mehrstufige Bewertungen wie z. B. „teils relevant“ sind aus Auswertungsgründen nicht möglich. Die grundlegenden Entscheidungskriterien für die Relevanz eines Dokumentes werden den Juroren vor der Bewertung genannt (siehe unten). Die gesamten Topic-Informationen (Kurz- und Langbeschreibung) sowie allgemeine Bewertungs-Richtlinien sollen von den Juroren in den Bewertungsprozess mit einbezogen werden. In dieser Reihenfolge geben folgende Dokumenteigenschaften den Ausschlag für die Bewertung:

1. Titel: Der Titel des Dokuments stimmt mit der Fragestellung überein.
2. Abstract (falls vorhanden): Das Abstract spiegelt die Fragestellung wider.
3. Schlagwörter: Die Schlagwörter lassen auf das Thema des Dokuments schließen.
4. Klassifikation: Die Benennung der Klassifikationsnotation lässt auf die Thematik des Dokuments schließen.

Das Erscheinungsjahr, der Dokumenttyp oder die Quelle (z. B. Zeitschrift oder Verlag) sind als weniger wichtig für die Bewertung zu erachten. In den IR-Tests konnten aufgrund der Sprachkenntnisse der Juroren lediglich Dokumente in Deutsch oder Englisch einbezogen werden. Dokumente in anderen Sprachen wurden als nicht-relevant bewertet.

Je nach Dokumentenmenge kann die Interpretation der Relevanz unterschiedlich weit ausgelegt werden, folgende Hinweise sollten allerdings beachtet werden:

- Stellt der Inhalt des Dokuments eine Teilmenge der Fragestellung dar (z. B. in einem Thema wird nach geographischem Umfeld Europa gesucht und das Dokument behandelt die Benelux-Staaten), wird das Dokument als relevant erachtet.
- Stellt der Inhalt des Dokumentes eine Obermenge der Fragestellung dar (z. B. in einem Thema wird nach geographischem Umfeld Europa gesucht und das Dokument behandelt die gesamte Welt), sollte das Dokument nur als relevant betrachtet werden, wenn die Fragestellung explizit erwähnt wird (also z. B. Europa, Asien etc.) und genügend Indizien bestehen, dass die Fragestellung hinreichend behandelt wird.

Wichtigstes Kriterium bei der Bewertung ist, dass die Juroren pro Topic und Dokumentpool konsistent bewerten, also den gleichen Typ Dokument entweder als relevant oder nicht-relevant einstufen.

Ein weiteres Kriterium bei der Evaluation von Literaturdokumenten ist die Auswahl der Juroren. Die Qualifikation und Ausbildung der Personen muss mit den zu bewertenden Fragestellungen und Inhalten der Datenbanken abgestimmt sein. Erfahrungen bei der inhaltlichen Erschließung (Indexierung) und/oder der Recherche im Informations- und Dokumentationsbereich sind von großem Vorteil für eine zügige und kompetente Bewertung der Dokumente.

Im CLEF-Kontext werden in der Regel 3 - 5 Personen für die Bewertung der unterschiedlichsprachigen Dokumente eingesetzt. Diese Personen haben meist langjährige

material that is appropriate to the requirement statement.” (Saracevic, Tefko (1975): *Relevance: A review of and a framework for the thinking on the notion in information science*, *Journal of the American Society for Information Science* 26 [6], Seite 321-343.)

Erfahrung in der Bewertung von Dokumenten und sind i. d. R. Muttersprachler in der Sprache der zu bewertenden Dokumente.

Im Projekt KoMoHe haben insgesamt 2 Personen die Dokumentbewertungen durchgeführt: eine sozial- und wirtschaftswissenschaftlich ausgebildete Person, die die Dokumente der Datenbanken aus dem Bereich Sozialwissenschaften, Wirtschaftswissenschaft und Internationale Beziehungen bewertet hat (siehe Übersicht in Tabelle 16), und eine Person mit Medizin-Ausbildung, die die Medline- und Psynindex-Dokumente aus dem Bereich der Medizin und Psychologie bewertet hat.

Im Projekt KoMoHe wurde ein eigenes Bewertungs-Tool eingesetzt, das von den Juroren installiert wird und alle Dokumente enthält, die bewertet werden sollen (siehe Abbildung 5-9). **Das Bewertungs-Tool wurde von S. Bärish (GESIS, Abteilung „Informationsprozesse in den Sozialwissenschaften“) in Java entwickelt und im Rahmen der Dokumentbewertungen der KoMoHe-Evaluation und CLEF eingesetzt.** Seit 2007 werden die Bewertungen des CLEF-Projekts zentral und datenbankgestützt durchgeführt (Nunzio und Ferro, 2005).

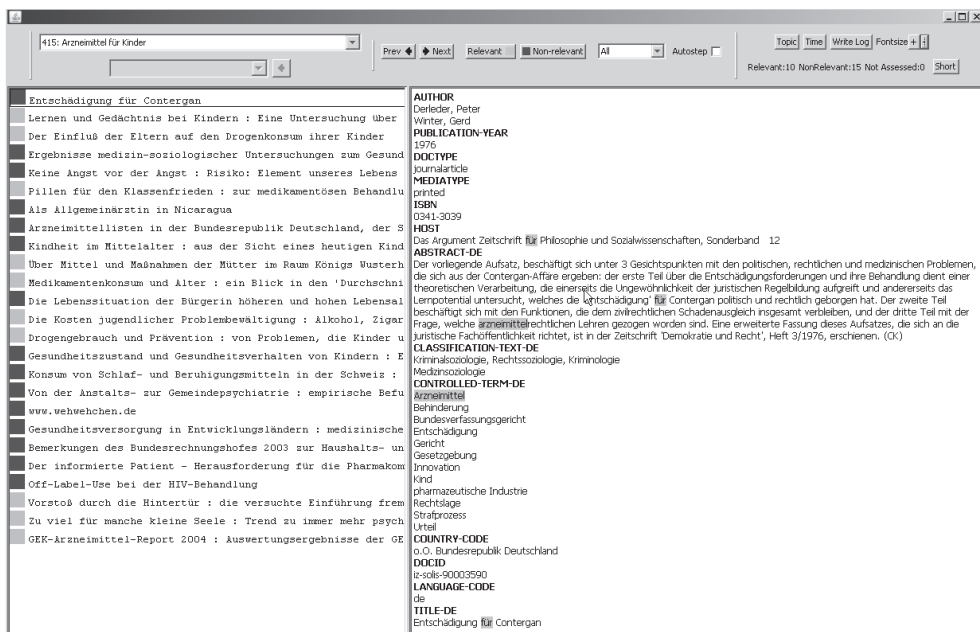


Abbildung 5-9: Screenshot des Assessment-Tools. Beispiel-Topic aus der KoMoHe-Evaluation (Nr. 415: Arzneimittel für Kinder)

Die Terme aus der Fragestellung werden in der Detailansicht rot markiert (siehe Abbildung 5-9). Folgende Funktionen stellt das Assessment-Tool den Juroren zur Verfügung:

1. Anzeige der Fragestellung (Topic in Kurzform und Langfassung)

2. Dokumente, die für Fragestellung auf Relevanz evaluiert werden sollen, sind über ihren Titel selektierbar (siehe Titeldarstellung, linke Bildschirmhälfte in Abbildung 5-9)
3. Detailansicht des selektierten Dokumentes, das auf Relevanz geprüft wird (siehe rechte Bildschirmhälfte in Abbildung 5-9)
4. Übersicht über als relevant / nicht-relevant schon geprüfte Dokumente sowie die Gesamtanzahl zu prüfender Dokumente (rote und grüne Markierungen (linker Rand) sowie Bewertungsstatistik im oberen Bildschirmabschnitt)
5. Erstellung der Logdatei mit den Relevanzbewertungen (siehe „Write Log“-Funktion)

```
10.2452/180-DS GIRT-DE19900100353 0
10.2452/180-DS GIRT-DE19900100966 1
10.2452/180-DS GIRT-DE19900100999 0
10.2452/180-DS GIRT-DE19900101529 1
```

Abbildung 5-10: Vier bewertete GIRT-Dokumente für das Topic 180 (fett). Die Relevanzurteile (fett) finden sich neben der Dokumenten-Nummer (rechts).

Die binären Relevanzurteile der Juroren werden bei der Bewertung mit den Dokumenten-IDs und den Topic-Nummern verknüpft und gespeichert (siehe Ausschnitt einer Logdatei nach Bewertung in Abbildung 5-10).

Nach Abschluss der Dokumentbewertungen können auf Basis der Logfile-Einträge die Auswertungen für jedes Topic erstellt werden (siehe Beispiel in Tabelle 24).

Die angewendeten Evaluationsmaße werden im folgenden Abschnitt vorgestellt.

5.7 Evaluationsmaße

Precision und Recall sind die beiden Standardmaße im Information Retrieval und werden als Effektivitäts- und Qualitätsmaße zur Untersuchung von Retrieval-Algorithmen häufig verwendet. Grundlage der Berechnung dieser beiden Maße ist der Anteil der relevant bewerteten Dokumente zu einer Suchanfrage (vgl. (van Rijsbergen, 1979)). Wie bereits im vorherigen Abschnitt beschrieben, wird für die Precision- und Recall-Berechnung vorausgesetzt, dass die Bewertungen der Relevanz der Suchergebnisse binär erfolgen.

Die Recall-Ratio oder kurz der Recall misst das Verhältnis von gefundenen relevanten Dokumenten zu einer Suchanfrage an allen relevanten Dokumenten aus einer Kollektion. Der Recall misst damit die Vollständigkeit eines Retrieval-Ergebnisses. Wenn alle relevanten Dokumente einer Kollektion zu einer Fragestellung gefunden werden, nimmt der Recall den Wert 1 an.

Die Precision misst das Verhältnis von gefundenen relevanten Dokumenten zu einer Suchanfrage an allen gefundenen Dokumenten aus einer Kollektion. Die Precision misst damit die Präzision bzw. Exaktheit eines Retrieval-Ergebnisses. Wenn nur relevante Dokumente aus der Kollektion zu einer Fragestellung gefunden werden, nimmt die Precision den Wert 1 an.

Die beiden Maße können formal folgendermaßen ausgedrückt werden:

Recall = relevante Dokumente \cap gefundene Dokumente / alle relevanten Dokumente

Precision = relevante Dokumente \cap gefundene Dokumente / alle gefundenen Dokumente

Es existieren weitere spezialisierte Maße im Information Retrieval, die aber hier nicht beschrieben werden sollen.

Für die Evaluation der Topics in dieser Studie wurden die beiden Standardmaße je Zone (Core, z2, z3) nach einem veränderten Verfahren berechnet.

5.7.1 Precision- und Recall-Berechnung

Um den Effekt des Re-Ranking (Bradfordizing) zu evaluieren, wurde vor allem das Maß Precision auf Basis der Relevanzurteile der gefundenen Dokumente je Zone errechnet und anschließend zonenweise verglichen. Da beim Re-Ranking in diesem Experiment an der Ergebnismenge keine Veränderungen vorgenommen werden, also keine neuen Dokumente hinzukommen, spielt das Maß Recall, das auf die Vollständigkeit eines Suchergebnisses fokussiert, eine untergeordnete Rolle. Beim Re-Ranking von Dokumenten sollen vielmehr die relevanten Dokumente einer Treffermenge in den vorderen Trefferlistenbereich gerankt werden. Ein Vergleich der Precision-Werte für unterschiedliche Trefferlistenbereiche (Dokument-Zonen) kann dies untersuchen.

Besonderheit der Berechnung ist die Bestimmung der Precision je Zone und zudem die separate Berechnung für die beiden evaluierten Dokumenttypen Zeitschriftenartikel und Monographien.

Die folgenden Maßzahlen wurden pro Topic analysiert:

- Retrieved: Anzahl der Dokumente im Dokumentenpool (pro Zone): Retrieved_article: Anzahl der Zeitschriftenartikel im Dokumentenpool (pro Zone) und Retrieved_book: Anzahl der Monographien im Dokumentenpool (pro Zone)
- Relevant: Anzahl der relevanten Dokumente (pro Zone und Dokumenttyp)
- Recall: Anteil der relevanten Dokumente an allen relevanten Dokumenten (Durchschnitt pro Zone)
- Precision: Anteil der relevanten gefundenen Dokumente an allen Dokumenten (Durchschnitt pro Zone)

Die beiden angepassten Maßzahlen Recall_z und Precision_z wurden zu diesem Zweck eingeführt:

Recall_z = relevante Dokumente pro Zone / alle relevanten Dokumente der Zonen

Precision_z = relevante Dokumente pro Zone / alle Dokumente pro Zone

Der pro Zone berechnete Recall (hier Recall_z genannt) misst das Verhältnis von relevanten Dokumenten zu einem Topic je Zone an allen relevanten Dokumenten aus dem bewerteten Dokumentenpool (Σ Zone 1 + Zone 2 + Zone 3). Es lassen sich folgende drei Recall-Werte pro Topic bestimmen:

- Recall_{z1} = Recall in Zone 1 oder Core
- Recall_{z2} = Recall in Zone 2

- $\text{Recall}_{z3} = \text{Recall in Zone 3}$

Die pro Zone berechnete Precision (hier Precision_z genannt) misst das Verhältnis von relevanten Dokumenten zu einem Topic je Zone an allen Dokumenten je Zone. Es lassen sich folgende drei Precision-Werte pro Topic bestimmen:

- $\text{Precision}_{z1} = \text{Precision in Zone 1 oder Core}$
- $\text{Precision}_{z2} = \text{Precision in Zone 2}$
- $\text{Precision}_{z3} = \text{Precision in Zone 3}$

Für das Experiment wurden die obigen Maßzahlen für jede Zone separat bestimmt und die Durchschnittswerte der einzelnen Cluster (core, z2, z3) verglichen.

Die Bewertungsdaten der beiden Evaluationsprojekte wurden nach dem gleichen Muster analysiert. Die ersten Relevanzanalysen wurden mit den KoMoHe-Datensätzen am Beispiel der 39 Topics durchgeführt. Die Daten der CLEF-Initiative sollen auf einer deutlich breiteren Basis (mehr Fragestellungen) die Effekte des Re-Rankings nach Bradfordizing verdeutlichen.

5.7.2 Beispielsberechnung

Das folgende Beispiel (Topic Nr. 171 „Computer im Alltag“, vgl. Tabelle 24) soll die vereinfachte Berechnung der Standardmaße Precision und Recall in unserer Analyse illustrieren. Die Auswertung der Relevanzbewertungen konzentriert sich vor allem auf die Bestimmung der Precision. Die Bestimmung des Recall ist in unserem Untersuchungsdesign aus mehrerlei Gründen weniger aussagekräftig und spielt in dieser Studie eine vernachlässigbare Rolle.

Beispiel: Der GIRT-Korpus (Datenbank SOLIS) enthält zum Zeitpunkt des Retrieval-Tests 208 Zeitschriftenartikel, die bezüglich der Relevanz zum Topic intellektuell bewertet wurden. 80 Dokumente wurden als relevant und 128 Dokumente wurden als nicht relevant für das Topic bewertet. Damit ergibt sich eine Precision_z von 0,38 für alle Zeitschriftenartikel (Berechnung: $\text{Relevant}/\text{Retrieved_article}$). Die Auswertung der separaten Precision_z für die drei Bradford-Zonen findet sich ebenfalls in Tabelle 24. Für die einzelnen Zonen wurde zusätzlich der Recall_z bestimmt (z. B. Berechnung des Recall_z für Zone 1: $\text{Relevant_Zone 1}/\text{Relevant } \Sigma \text{ Zone 1 + Zone 2 + Zone 3}$).

Auf Basis der Precision_z -Werte lassen sich die prozentuellen Precision_z -Unterschiede zwischen den Zonen errechnen. Im Beispiel ist z. B. die Precision der Zone 1 um 180 % höher als die Precision der Zone 3.

Berechnung: $(\text{Precision}_{z1}/\text{Precision}_{z3}) * 100 - 100$

Tabelle 24: Beispiel Precision- und Recall-Auswertung für Zeitschriftenartikel des Topics Nr. 171

| | Retrieved_article | Relevant | Precision _z | Recall _z |
|----------|-------------------|----------|------------------------|---------------------|
| Σ Zone 1 | | | | |
| + Zone 2 | | | | |
| + Zone 3 | 208 | 80 | 0,38 | k. A. |
| Zone 1 | 73 | 41 | 0,56 | 0,51 |
| Zone 2 | 65 | 25 | 0,38 | 0,31 |
| Zone 3 | 70 | 14 | 0,20 | 0,18 |

Die Ergebnisse der Precision-Berechnungen finden sich in Abschnitt 7.4. Die Recall-Werte werden in dieser Studie nicht näher untersucht.

5.8 Signifikanztest

Im Anschluss an die Berechnung der oben genannten Maßzahlen wurden die durchschnittlichen Precision_z-Werte pro Zone auf signifikante Unterschiede hin untersucht. Es wurden folgende Signifikanztests (vgl. (Bortz, 2005)) durchgeführt⁷¹:

- **T-Test bei gepaarten Stichproben:** Beim T-Test bei gepaarten Stichproben werden die Mittelwerte zweier Variablen miteinander verglichen. Dieser Test berechnet für jeden Fall die Differenzen zwischen den Werten der Variablen und überprüft, ob der Durchschnitt von Null abweicht. Der T-Test ist ein Signifikanztest für kontinuierlich skalierte und (zumindest annähernd) normalverteilte Variablen. Es werden zwei Mittelwerte verglichen, die aus einer Stichprobe stammen; so wird der T-Test auf abhängige (oder gepaarte) Stichproben angewandt. Der T-Test berücksichtigt die Verteilung der Mittelwerte von Differenzen zwischen den beiden Variablen.
- **Wilcoxon-Test bei zwei verbundenen Stichproben:** Der Wilcoxon-Test ist ein nicht-parametrischer Test, der prüft, ob die Verteilung von zwei gepaarten Variablen in zwei verbundenen Stichproben gleich ist. Der Test berücksichtigt die Größe der Unterschiede zwischen den beiden gepaarten Variablen. Der Wilcoxon-Test ist ebenfalls ein Signifikanztest, der ein robusteres Maß (relativ zum T-Test) darstellt. Er wird üblicherweise bei ordinalem Skalenniveau verwandt oder im Falle, dass die Mittelwerte negative Vorzeichen haben.

Die beiden Tests wurden a) jeweils für die Gesamtheit der Topics einer Evaluationsreihe (CLEF bzw. KoMoHe) und b) für Untergruppen der jeweiligen Evaluationsreihe (z. B. 25 Topics des CLEF-Jahrgangs 2003) durchgeführt. Die Tests wurden sowohl für Zeitschriftenartikel als auch für die analysierten Monographien eingesetzt.

⁷¹ Alle Signifikanztests wurden mit der Software SPSS in der Version 14 durchgeführt.

Die Signifikanztests prüfen paarweise auf signifikante Unterschiede zwischen:

- $Precision_z$ im Core und $Precision_z$ in Zone 3 ($Precision_{z1} - Precision_{z3}$)
- $Precision_z$ im Core und $Precision_z$ in Zone 2 ($Precision_{z1} - Precision_{z2}$)
- $Precision_z$ in Zone 2 und $Precision_z$ in Zone 3 ($Precision_{z2} - Precision_{z3}$)
- $Precision_z$ im Core und $Precision_z$ für alle Zonen ($Precision_{z1} - Precision_z$)

Die Ergebnisse der Signifikanztests finden sich in Kapitel 7. Die vollständigen Signifikanztests finden sich im Anhang „Signifikanztests“.

5.9 Qualitatives Kurzinterview

Die Anwendung informatrischer und vergleichbarer Verfahren ergibt eine alternative Sicht auf Ergebnismengen, die vom Benutzer i. d. R. als intuitiv und positiv empfunden wird (vgl. z. B. Autorenzentralität bei (Mutschke, 2003, Mutschke, 2004a, Mutschke, 2004b) oder die Verfahren PageRank (Chen et al., 2007, Page et al., 1998) oder der Kleinberg-Algorithmus (Kleinberg, 1999)).

„Autorennetzwerk-Retrievalmodelle bieten also offensichtlich eine ganz andere Sicht auf eine Dokumentkollektion als traditionelle termbasierte Verfahren, deren Ergebnismengen sich oftmals erheblich überlappen. Dieses Ergebnis zeigt, dass Retrievalmodelle, die auf Zentralität in Autorennetzwerken basieren, einen erheblichen informationellen Mehrwert haben können.“ ((Mutschke, 2004) S. 45)

Autorennetzwerke (bzw. Ko-Autorennetzwerke) sind schon seit längerem Gegenstand scientometrischer und netzwerkanalytischer Untersuchungen (vgl. (Wasserman und Faust, 1994)). Empirische Studien zeigen, dass wissenschaftliche Kooperationsnetzwerke in der Regel ein bestimmtes Organisationsmuster aufweisen, nämlich eine ausgeprägte Small-World-Architektur mit einer starken Tendenz zur Zentralisierung (vgl. (Mutschke, 2004)). Es liegt daher nahe, Wissen über die Struktur von Kooperationsnetzwerken und die strategische Position der Akteure in diesen Strukturen auch für informationswissenschaftliche Kontexte wie Information Retrieval und Datenbankanalyse zu nutzen. Netzwerkanalytische Konzepte werden bisher jedoch nur in sehr eingeschränktem Maße in informationswissenschaftlichen Domänen eingesetzt.

Das Bradfordizing einer Ergebnismenge liefert ebenfalls eine alternative Sicht auf ein beliebiges Suchergebnis mit vergleichbaren Effekten.

Zum Abschluss des Kapitels wird der Aufbau eines qualitativen Interviews beschrieben, das parallel zu den bisher beschriebenen Analysen durchgeführt wurde und Kernzeitschriften zu unterschiedlichen Themengebieten thematisiert.

Der domain-specific Track der CLEF-Initiative setzt sich jedes Jahr aus 25 unterschiedlichen Themen vor allem aus dem sozialwissenschaftlichen Themenspektrum zusammen (siehe dazu Abschnitt 5.4.1). Im Vorfeld der jährlichen Retrieval-Konferenz wurden in den vergangenen Jahren z. B. von den Mitarbeiterinnen und Mitarbeitern der GESIS-Fachinformationsabteilung Vorschläge für Themen (im CLEF-Kontext Topics genannt) zur Retrieval-Evaluation erbeten. Aus den Vorschlägen werden 25 Topics ausgewählt und für die weitere Evaluation aufbereitet (siehe dazu (Kluck, 2002, Petras et al., 2007)).

Die untersuchten Topics des Jahres 2007 bzw. die Personen, die diese Topics vorgeschlagen haben, stehen im Mittelpunkt des qualitativen Kurzinterviews.

Im Rahmen dieser Studie wurden 12 Personen⁷² der GESIS durch den Autor in einem strukturierten (geleiteten) Kurzinterview befragt. Diese Personen wurden per Email und Telefon kontaktiert und gebeten, an dem Kurzinterview teilzunehmen; anschließend wurden Termine für die Kurzbefragung vereinbart. Insgesamt konnten 24 Topics in die Befragung aufgenommen werden⁷³. Die Befragung der Personen wurde im Februar 2008 durchgeführt. Damit liegt der Befragungstermin relativ zeitnah zu der Einreichung der Vorschläge (etwa ein Jahr nach der Einreichung). Das Interview besteht aus vier kurzen Fragen, die teils per Telefon, teils in „Face-to-Face“-Gesprächen mit den 12 beteiligten Personen gestellt wurden. Bevor die erste Frage gestellt wurde, wurde den Interviewpartnern erklärt, was eine Kernzeitschrift ausmacht und wie diese für das Interview analysiert wurden (quantitative Auszählung, vgl. das Verfahren Bradfordizing). Den meisten Personen war der Begriff Kernzeitschrift ohnehin bekannt.

Grundsätzlich muss einschränkend zum Kurzinterview bemerkt werden:

- Die befragte Gruppe besteht aus wenigen Teilnehmern und ist damit nicht repräsentativ.
- Wie viele qualitativen Techniken unterliegen die Ergebnisse der Auswertung und Interpretation.
- Der Interviewer (in diesem Fall der Autor) nimmt eine zentrale moderierende Rolle ein und kann daher starken, u. U. auch negativen Einfluss auf Verlauf und Inhalt des Interviews ausüben.

Folgende vier Fragen wurden den Personen gestellt:

- Frage 1: Nennen Sie mir bitte spontan mehrere Zeitschriften, die Sie zu den Kernzeitschriften (häufigsten Zeitschriften) des Themas (XY⁷⁴) in der Datenbank SOLIS zählen.
- Frage 2⁷⁵: Hätten Sie diese Zeitschriften als häufigste Veröffentlichungsorgane für Ihr Topic erwartet?
- Frage 3: Können Sie sich vorstellen, dass die eben vorgelesenen Kernzeitschriften relevantere Arbeiten zum Thema publizieren als die übrigen in SOLIS ausgewerteten Zeitschriften?
- Frage 4: Welche Zeitschriften, die regelmäßig Literatur zu den obigen Topics beinhalten, vermissen Sie in dieser Liste? Zählen Sie bitte mehrere Zeitschriftentitel auf!

Ein vollständiger Beispielfragebogen findet sich im Anhang (siehe Anhang „Beispielfragebogen“).

72 Der Autor dankt allen Mitarbeiterinnen und Mitarbeitern, die an der Befragung teilgenommen haben.

73 Ein Topic konnte nicht zugeordnet werden und wurde deshalb bei der Befragung nicht berücksichtigt.

74 Hier wurde der ursprüngliche Titel des Themas eingesetzt.

75 Bevor Frage 2 gestellt wurde, wurden die analysierten Kernzeitschriften des Themas vorgelesen.

Im Mittelpunkt des Kurzinterviews standen inhaltlich nicht die vorgeschlagenen Topics, sondern die spezifische Zeitschriftenliteratur zu den jeweiligen Fragestellungen in der Datenbank SOLIS. Die befragten Personen waren zum Zeitpunkt der Interviews entweder konkret mit dem Datenbankaufbau von SOLIS beschäftigt oder zumindest mit der Recherche in sozialwissenschaftlichen Datenbanken (z. B. SOLIS) vertraut. Es konnte daher davon ausgegangen werden, dass den befragten Personen die einschlägige Literatur in den deutschsprachigen Sozialwissenschaften zu den konkreten Fragestellungen bekannt war. Des Weiteren muss erwähnt werden, dass die Interviewpartner auf sehr unterschiedliche Erfahrungshorizonte (einjährige vs. jahrzehntelange Erfahrung) beim Aufbau der Literaturbestände der Datenbank SOLIS zurückgreifen können.

Hintergrund der Interviews war vor allem der Wunsch nach mehr qualitativen Einschätzungen zur Bedeutung der Kernzeitschriften für spezifische Themen. Bei der Konzeption des Fragebogens war von vornherein klar, dass die wenigen nichtrepräsentativen Interviews nicht dazu geeignet sind, Thesen zu verifizieren oder zu falsifizieren, vielmehr sollten die individuellen Kenntnisstände und Erfahrungen auf der Mitarbeiterebene erhoben, punktuell analysiert und anschließend in Bezug auf die Thesen dieser Arbeit diskutiert werden. Die Fragen sollen den Interpretationsspielraum der zuvor dargestellten quantitativen Ergebnisse somit erweitern und u. U. Thesen in Teilaussagen untermauern. Die Ergebnisse der Befragung finden sich in Abschnitt 7.3.

Im folgenden Kapitel werden einzelne Vorstudien beschrieben, die explorative Vorarbeiten zur Durchführung der Hauptuntersuchung (siehe Kapitel 5.2) darstellen. Die Vorstudien erkunden Fragestellungen, die nicht integraler Bestandteil der Hauptuntersuchung sind, die aber Hinweise und Anschlussmöglichkeiten der in dieser Studie vorgestellten Verfahren liefern.

6 Effekte von Crosskonkordanzen und Bradfordizing

Im Mittelpunkt der Vorstudien stehen a) die Effekte der Verfahren zur Heterogenitätsbehandlung (insbesondere Crosskonkordanzen) und b) die Beobachtung informatrischer Gesetzmäßigkeiten (insbesondere Bradford's Law of Scattering bzw. Power Laws) in Literaturdatenbanken und vergleichbaren Dokumentkollektionen. Insgesamt werden fünf Untersuchungen vorgestellt, die die Mehrwerte und Auswirkungen der in dieser Arbeit untersuchten Dienste verdeutlichen.

- In Abschnitt 6.1 werden die Auswirkungen des Einsatzes von Crosskonkordanzen am Beispiel von Themen in den Sozialwissenschaften beschrieben. Diese Vorstudie ergänzt die Ergebnisse der Evaluation der Crosskonkordanzen in Abschnitt 5.2.1.
- Abschnitt 6.2 demonstriert die Allgemeingültigkeit des Verfahrens Bradfordizing (siehe Abschnitt 5.1) für Ein-Wort-Anfragen in unterschiedlichen Literaturdatenbanken.
- Abschnitt 6.3 wendet das Verfahren Bradfordizing auf die Deskriptor-Anfragen einzelner soFid-Themen aus Abschnitt 6.1 an.
- Die Vorstudie in Abschnitt 6.4 untersucht die Literatur in der Datenbank LISA und beschreibt einen Mehrwert des Verfahrens Bradfordizing.
- Abschnitt 6.5 geht der Frage nach, inwiefern sich die Zeitschriften in den drei Zonen (core, z2 und z3) bzgl. ihrer Erscheinungsweise (Periodizität) unterscheiden.
- In Abschnitt 0 wird versucht, Bradfordizing auf Internet-Dokumente anzuwenden.

Die explorativen Vorstudien wurden alle im Rahmen des Projekts KoMoHe durchgeführt und sind bereits z. T. in Einzelpublikationen veröffentlicht worden.

6.1 Überlappung und Termzugewinn durch Crosskonkordanzen („soFid-Experiment“)

In Ergänzung zu der themenzentrierten Evaluation im Projekt KoMoHe (vgl. Abschnitt 5.2.1), bei der Fragestellungen durch wenige Suchbegriffe operationalisiert wurden (zwischen 1 - 3 Termen bei der Freitextsuche und 2 - 6 Termen bei der Schlagwortsuche), fokussiert das „soFid⁷⁶-Experiment“ auf die Operationalisierung von thematischen Suchterm-Gruppen. Das Experiment soll vor allem die Auswirkungen breiter Deskriptoren-Kombinationen und deren Überlappungseffekte beim Einsatz von Crosskonkordanzen am Beispiel von deutsch- und englischsprachigen Datenbanken demonstrieren. In einem Informationsportal mit heterogenen Datenbanken ist insbesondere entscheidend, ob eine Crosskonkordanz eine verteilte Suche ermöglichen kann. Können Crosskonkordanzen z. B. die Sprachbarriere überbrücken, um eine nahtlose Suche mit der gleichen Anfrage in unterschiedlichen Datenbanken zu ermöglichen?

76 soFid steht für sozialwissenschaftlicher Fachinformationsdienst. Die soFids stellen den aktuellen Stand der Literatur und Forschung des jeweiligen Gebiets aus den Datenbanken SOLIS und SOFIS zusammen. Siehe dazu <http://www.gesis.org/dienstleistungen/fachinformationen/servicepublikationen/sofid/>

Jedes der folgenden Themen wird in der Praxis durch eine Reihe von Deskriptoren, die logisch über boolesche OR-Verknüpfungen verbunden werden, unterschiedlich breit operationalisiert (vgl. Tabelle 25) und in unterschiedlichen Datenbanken abgefragt. In das Experiment wurden insgesamt 11 soFid-Themen und sieben Datenbanken des Suchportals sowiport sowie ireon⁷⁷ einbezogen. Die Themen stammen alle aus den Sozialwissenschaften bzw. einzelnen sog. Bindestrich-Soziologien.

Tabelle 25: soFid-Themen und die Anzahl der Deskriptoren des Thesaurus Sozialwissenschaften (TheSoz) pro Thema

| soFid-Thema | Operationalisierung (Anzahl der Deskriptoren) |
|--|--|
| Allgemeine Soziologie | 14 |
| Bevölkerungsforschung | 20 |
| Gesellschaftlicher Wandel in den neuen Bundesländern | 7 |
| Gesundheitsforschung | 54 |
| Industrie- und Betriebssoziologie | 17 |
| Kriminal-, Rechtssoziologie | 17 |
| Religionsforschung | 23 |
| Umweltforschung | 51 |
| Wissenschafts- und Technikforschung 1 | 10 |
| Wissenschafts- und Technikforschung 2 | 29 |
| Wissenschafts- und Technikforschung 3 | 33 |

Für das Experiment wurden die Deskriptoren der einzelnen Themen durch den in Abschnitt 3.2.5 beschriebenen Transformations-Service (HTS) in die entsprechenden Vokabulare der folgenden Datenbanken transformiert und mit den ursprünglichen Deskriptoren verglichen. Es wurden für das Experiment ausschließlich äquivalente Term-Transformationen verwendet.

Das folgende Beispiel zeigt zwei Deskriptoren des soFid-Themas „Bevölkerungsforschung“ und die Transformationen des kontrollierten Vokabulars der Datenbank ASSIA zu diesen Deskriptoren. Der Deskriptor „Altersteilzeit“ wird in die beiden AND-verknüpften Begriffe „Elderly people“ und „Part time employment“ transformiert. Die zwei deutschen Terme ergeben damit insgesamt drei neue Begriffe. Diese neuen Terme sollen im Folgenden als Zugewinn bezeichnet werden (vgl. Tabelle 26 und Tabelle 27). **Der Zugewinn an Termen wäre in diesem Fall drei.**

TheSoz: Geburtenfolge OR Altersteilzeit

ASSIA: „Birth Order“ OR („Elderly people“ AND „Part time employment“)

⁷⁷ <http://www.ireon-portal.de>

Folgende Datenbanken wurden für die Untersuchung berücksichtigt:

- Solit: erfasst Literatur zur Sozialarbeit und Sozialpädagogik. Der zugrunde liegende Thesaurus und die Datenbank sind deutschsprachig.
- Gerolit: erfasst Literatur zur sozialen und psychologischen Gerontologie und Altenarbeit. Der zugrunde liegende Thesaurus und die Datenbank sind deutschsprachig.
- WAO: World Affairs Online erfasst Literatur zu Internationaler Politik, Internationaler Sicherheit, Weltwirtschaft, Internationalem System und Länder- und Regionalforschung. Der zugrunde liegende Thesaurus ist deutschsprachig. WAO wird z.B. im Suchportal ireon angeboten.
- WPSA: Worldwide Political Science Abstracts erfasst Literatur aus den Politikwissenschaften. Der zugrunde liegende Thesaurus und die Datenbank sind englischsprachig.
- PEI: Physical Education Index erfasst Literatur zu den Themen Sportsoziologie, Sportpsychologie sowie allgemein zu Sport und Gesundheit. Der zugrunde liegende Thesaurus und die Datenbank sind englischsprachig.
- CSA-SA: Sociological Abstracts erfasst Literatur aus den gesamten Sozialwissenschaften. Der zugrunde liegende Thesaurus und die Datenbank sind englischsprachig.
- ASSIA: Applied Social Sciences Index and Abstracts erfasst Literatur zu den Themen Gesundheits- und Sozialwesen, Soziologie, Kriminologie, Anthropologie, Bildung, Psychologie, Wirtschaft, Politikwissenschaft und Erziehung. Der zugrunde liegende Thesaurus und die Datenbank sind englischsprachig.

Tabelle 26 und Tabelle 27 zeigen die Anzahl der äquivalent transformierten Terme (siehe Transformation (ÄQ)) der entsprechenden Datenbanken. Zusätzlich wurden die transformierten Terme gegen die ursprüngliche Liste der Deskriptoren abgeglichen. Der Zugewinn an „neuen“ (nicht zeichengleichen) Termen durch die Transformationen wird ebenfalls pro Datenbank aufgeführt (siehe Zugewinn in Tabelle 26). Der Durchschnitt der Deskriptorzahlen über alle Themen wird in der letzten Zeile dargestellt.

Beispiel (Tabelle 26): Die 7 Deskriptoren (Thesaurus Sozialwissenschaften) des soFid-Themas „Gesellschaftlicher Wandel in den neuen Bundesländern“ können in 7 Deskriptoren des Solit-Vokabulars äquivalent transformiert werden (ÄQ-Relationen). 1 Solit-Term wird durch die Transformationen neu hinzugewonnen (vgl. Werte der Spalte 2 und 3 in Tabelle 26), d. h., dass die Transformationen in die Solit-Terme 6 zeichengleiche Terme liefern. Die Transformation in das Gerolit-Vokabular kann keine neuen Terme generieren, d. h. alle Terme kommen bereits im Thesaurus Sozialwissenschaften vor (vgl. Werte der Spalte 4 in Tabelle 26).

Tabelle 26: Äquivalente Transformationen und Deskriptoren-Zugewinn in deutschsprachigen Datenbanken

| soFid-Themen | Deskriptoren TheSoz | Transformation (ÄQ) Solit | Zugewinn Solit | Transformation (ÄQ) Gerolit | Zugewinn Gerolit | Transformation (ÄQ) WAO | Zugewinn WAO |
|---|------------------------|------------------------------|-------------------|--------------------------------|---------------------|----------------------------|-----------------|
| Allgemeine Soziologie | 14 | 8 | 6 | 5 | 4 | 9 | 4 |
| Bevölkerungsforschung | 20 | 17 | 13 | 14 | 6 | 11 | 8 |
| Gesellschaftlicher Wandel in den neuen Bundesländern | 7 | 7 | 1 | 7 | 0 | 7 | 2 |
| Gesundheitsforschung | 54 | 36 | 11 | 37 | 14 | 19 | 10 |
| Industrie- und Betriebssoziologie | 17 | 15 | 9 | 6 | 3 | 6 | 1 |
| Kriminal-, Rechtssoziologie | 17 | 13 | 9 | 7 | 4 | 14 | 4 |
| Religionsforschung | 23 | 14 | 12 | 8 | 6 | 13 | 3 |
| Umweltforschung | 51 | 27 | 26 | 30 | 25 | 41 | 20 |
| Wissenschafts- und Technikforschung 1 | 10 | 2 | 1 | 1 | 0 | 5 | 1 |
| Wissenschafts- und Technikforschung 2 | 29 | 17 | 5 | 19 | 6 | 26 | 7 |
| Wissenschafts- und Technikforschung 3 | 33 | 12 | 11 | 11 | 7 | 17 | 4 |
| Durchschnitt | 25 | 15,3 | 9,5 | 13,2 | 6,8 | 15,3 | 5,8 |

Tabelle 27 zeigt die Anzahl der äquivalent transformierten Terme (siehe Transformation (ÄQ)) der englischsprachigen Vokabulare/Datenbanken. Analog zu Tabelle 26 wird hier angegeben, wie viele äquivalente Deskriptoren aus dem Zielvokabular nach der Transformation durch den HTS generiert werden können. Die nächste Spalte gibt an, wie hoch der Zugewinn von neuen Deskriptoren ist, nachdem die Deskriptoren mit dem Ausgangsvokabular verglichen wurden.

Beispiel (Tabelle 27): Die 14 äquivalent transformierten Deskriptoren (CSA-SA-Thesaurus) des soFid-Themas „Industrie- und Betriebssoziologie“ führen aufgrund der Mehrsprachigkeit der Crosskonkordanz (Deutsch-Englisch) zu 14 neuen Deskriptoren in diesem Vokabular. Der gleiche Effekt trifft auch für die weiteren englischsprachigen kontrollierten Vokabulare zu (vgl. Werte in Fettdruck der Tabelle 27).

Tabelle 27: Äquivalente Transformationen und Deskriptoren-Zugewinn in englischsprachigen Datenbanken

| soFid-Themen | Transformation (ÄQ) CSA-SA | Zugewinn CSA-SA | Transformation (ÄQ) WPSA | Zugewinn WPSA | Transformation (ÄQ) ASSIA | Zugewinn ASSIA | Transformation (ÄQ) PEI | Zugewinn PEI |
|---|-------------------------------|--------------------|-----------------------------|------------------|------------------------------|-------------------|----------------------------|-----------------|
| Allgemeine Soziologie | 10 | 8 | 9 | 8 | 12 | 10 | 2 | 2 |
| Bevölkerungsforschung | 12 | 12 | 9 | 9 | 16 | 16 | 1 | 1 |
| Gesellschaftlicher Wandel in den neuen Bundesländern | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| Gesundheitsforschung | 21 | 21 | 20 | 20 | 40 | 40 | 10 | 10 |
| Industrie- und Betriebssoziologie | 14 | 14 | 9 | 9 | 11 | 11 | 2 | 2 |
| Kriminal-, Rechtssoziologie | 7 | 7 | 13 | 13 | 14 | 14 | 3 | 3 |
| Religionsforschung | 14 | 14 | 13 | 13 | 11 | 11 | 8 | 7 |
| Umweltforschung | 22 | 22 | 22 | 22 | 30 | 30 | 15 | 15 |
| Wissenschafts- und Technikforschung 1 | 4 | 4 | 5 | 5 | 7 | 7 | 0 | 0 |
| Wissenschafts- und Technikforschung 2 | 25 | 25 | 22 | 22 | 22 | 22 | 13 | 13 |
| Wissenschafts- und Technikforschung 3 | 21 | 21 | 19 | 19 | 16 | 16 | 6 | 6 |
| Durchschnitt | 13,7 | 13,5 | 12,8 | 12,7 | 16,4 | 16,2 | 5,5 | 5,4 |

Im Durchschnitt wurden 25 Deskriptoren pro soFid-Thema verwendet (vgl. Tabelle 26). Wie zu erwarten, zeigt die Analyse der Term-Transformationen der einzelnen Vokabulare, dass der Zugewinn an neuen Suchtermen bei den deutschsprachigen Datenbanken bzw. kontrollierten Vokabularen deutlich niedriger liegt als bei den englischsprachigen Vokabularen. Die transformierten Terme der englischsprachigen Vokabulare sind i. d. R. auch „neue“ Terme (vgl. Zugewinn in Tabelle 27). Der Anteil der Zeichengleichheit (Nulltransformationen bei Marx, 2005) bei den deutschsprachigen Vokabularen ist, wie zu erwarten, mit ca. 50 % zeichengleichen Terme hoch. In beiden Sprachbereichen können in etwa 50 % der ursprünglichen Deskriptoren (Thesaurus Sozialwissenschaften) in einen passenden Term im Zielvokabular überführt werden.

Das Experiment verdeutlicht zum einen die sprachliche Übersetzungsleistung der Crosskonkordanzen, zum anderen kann gezeigt werden, dass die Crosskonkordanzen selbst innerhalb einer Sprache und Disziplin die Anzahl der neuen kontrollierten Terme nahezu verdoppeln können. Der Einsatz von Crosskonkordanzen mit der in diesem Experiment demonstrierten Generierung neuer kontrollierter Suchterme zeigt damit positive Effekte für das Information Retrieval, insbesondere für die sprachübergreifende Suche.

Crosskonkordanzen ermöglichen dem Nutzer durch das Generieren neuer Suchterme eine Erweiterung der Suchmöglichkeiten und können damit die Informationssuche verbessern. Die Verbesserung der Suchergebnisse steht im Mittelpunkt eines weiteren Experiments dieser Studie (siehe Auswertungen zu Recall und Precision in Abschnitt 7.1).

6.2 Bradfordizing bei Ein-Wort-Anfragen

Im folgenden Abschnitt wird untersucht, ob sich die Bradford-typischen Verteilungen zu beliebigen Ein-Wort-Anfragen beobachten lassen. Ziel der Untersuchung ist, das generelle Auftreten der Bradfordschen Gesetzmäßigkeit für Zeitschriftenartikel in Fachdatenbanken, selbst für zufällige Ein-Wort-Anfragen, in einem einfachen Test zu verdeutlichen.

Für dieses Experiment wurden insgesamt 12 allgemeine Suchbegriffe aus dem Logfile des Suchportals sowiport ausgewählt. Die Terme wurden anschließend für die Suche in der englischsprachigen Datenbank (CSA-SA) im Test ins Englische übersetzt. In Tabelle 28 sind die deutschen Terme sowie die englische Übersetzung der Terme aufgeführt.

Tabelle 28: 12 deutsche Suchbegriffe mit der englischen Übersetzung

| deutsche Terme (Logfile) | englische Übersetzung |
|--------------------------|-----------------------|
| Alter | age |
| Armut | poverty |
| Bewohner | resident |
| Erfolg | outcome |
| Innovation | innovation |
| Jugend | adolescence |
| Karriere | career |
| Motivation | motivation |
| Pflege | care |
| Test | test |
| Verlauf | course |
| Werbung | advertising |

Die Vorstudie gliedert sich in folgende Schritte:

- Die 12 Terme wurden jeweils in Deutsch (SOLIS, WAO) und Englisch (CSA-SA) in den Datenbanken abgefragt.
- Die ersten 300 Zeitschriftenartikel (Metadaten der Dokumente) wurden pro Term und Datenbank heruntergeladen.
- Die ISSN-Nummern wurden aus den Metadaten der Zeitschriftenartikel extrahiert und nach dem Verfahren Bradfordizing (vgl. Abschnitt 5.1) analysiert. Hierzu wurde die Häufigkeit der einzelnen ISSN-Nummern ausgezählt und anschließend die Häufigkeit

der Artikel der einzelnen ISSN-Nummern kumuliert. Anschließend wurden die Verteilungen der Zeitschriftenartikel pro Term und Datenbank aufgetragen.

Die Abbildung 6-1, Abbildung 6-2 und Abbildung 6-3 zeigen die typischen Bradford-Verteilungen der Zeitschriftenartikel in den drei analysierten Datenbanken (SOLIS, WAO, CSA-SA). Exemplarisch wird jeweils zu einem Graphen eine Trendlinie hinzugefügt.

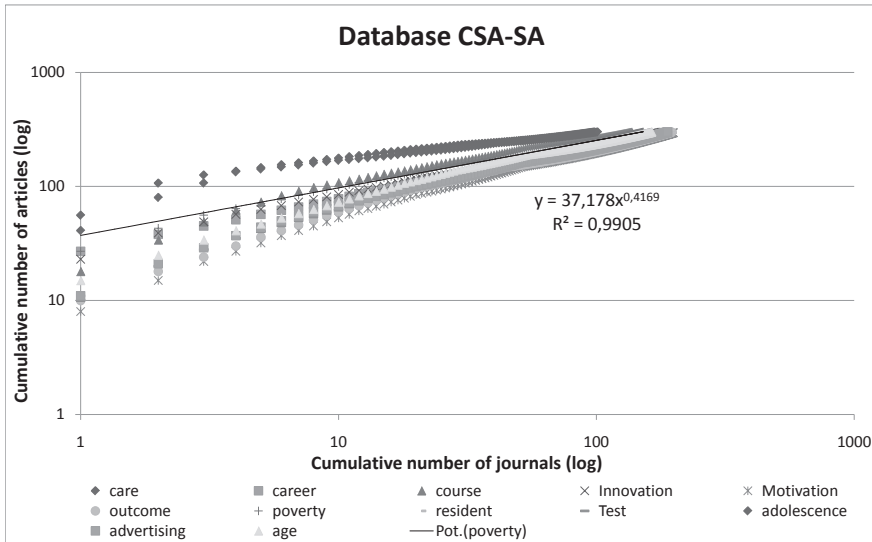


Abbildung 6-1: Die Verteilung der Zeitschriftenartikel für die englischsprachigen Terme in der Datenbank CSA-SA (doppelt-logarithmische Auftragung)

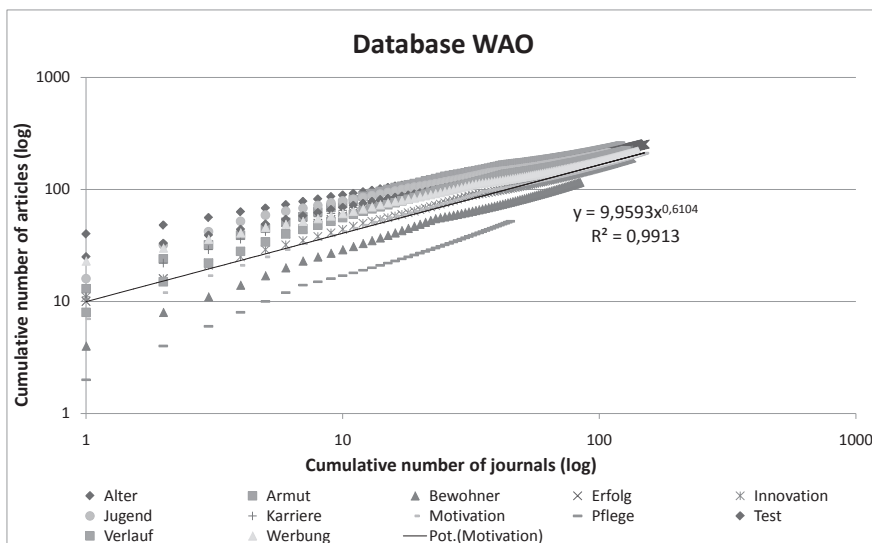


Abbildung 6-2: Die Verteilung der Zeitschriftenartikel für die deutschsprachigen Terme in der Datenbank WAO (doppelt-logarithmische Auftragung)

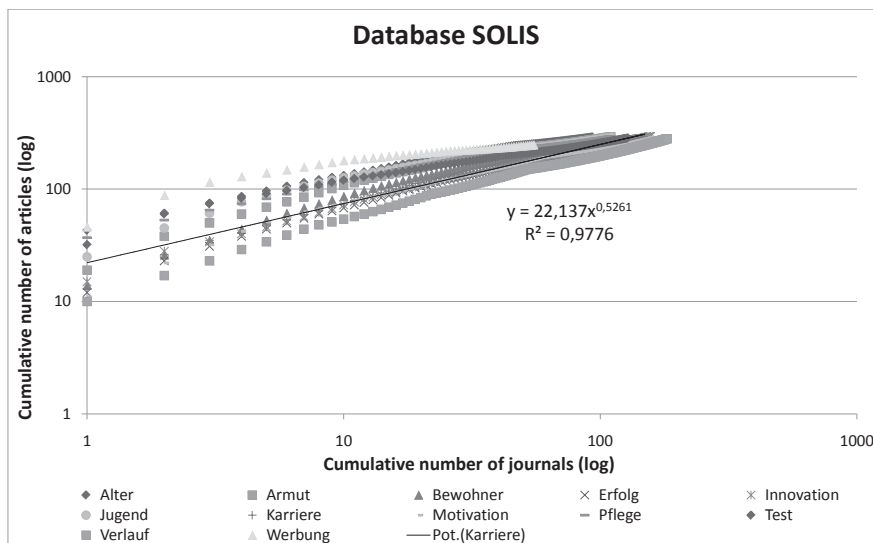


Abbildung 6-3: Die Verteilung der Zeitschriftenartikel für die deutschsprachigen Terme in der Datenbank SOLIS (doppelt-logarithmische Auftragung)

Die Analyse verdeutlicht, dass zu beliebigen allgemeinen Suchbegriffen die typischen Bradfordschen Verteilungen beim Dokumenttyp Zeitschriftenartikel auftreten. Prinzipiell kann das Verfahren Bradfordizing somit für allgemeine Ein-Wort-Anfragen angewendet werden. Die typischen Verteilungen lassen sich für alle Datenbanken bei der Analyse der ersten 300 Zeitschriftenartikel pro Anfrageterm beobachten. Dieses Ergebnis illustriert die Allgemeingültigkeit und damit Generalisierbarkeit dieses Verfahrens für Zeitschriftenartikel in Fachdatenbanken.

6.3 Bradfordizing bei Deskriptoren-Kombinationen

Im folgenden Abschnitt wird untersucht, ob sich die Bradford-typischen Verteilungen zu drei in Abschnitt 6.1 beschriebenen soFid-Themen (Deskriptoren-Kombinationen) beobachten lassen. Im Unterschied zu der vorherigen Studie in Abschnitt 6.2 (Ein-Wort-Anfragen) wurden anstatt der einzelnen Term-Anfragen thematisch spezifizierte Deskriptoren-Kombinationen als Anfragen an die Datenbanken geschickt und die zurückgelieferten Zeitschriftenartikel analysiert. Es wurden wie in Abschnitt 6.2 lediglich die ersten 300 Zeitschriftenartikel berücksichtigt. Die weiteren Schritte dieses Tests wurden analog durchgeführt.

Die Abbildung 6-4, Abbildung 6-5 und Abbildung 6-6 zeigen ebenfalls die typischen Bradford-Verteilungen der Zeitschriftenartikel zu drei soFid-Themen (vgl. Tabelle 26) in den analysierten Datenbanken (SOLIS, IBLK, WPSA, CSA-SA, PEI, ASSIA).

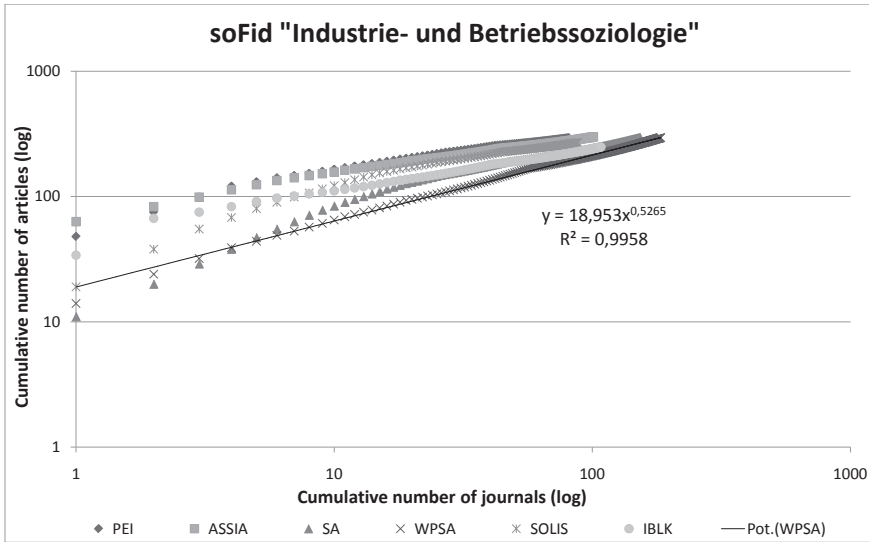


Abbildung 6-4: Die Verteilung der Zeitschriftenartikel für die Deskriptoren-Anfrage des soFid-Themas „Industrie- und Betriebssoziologie“ in den Datenbanken (doppelt-logarithmische Auftragung)

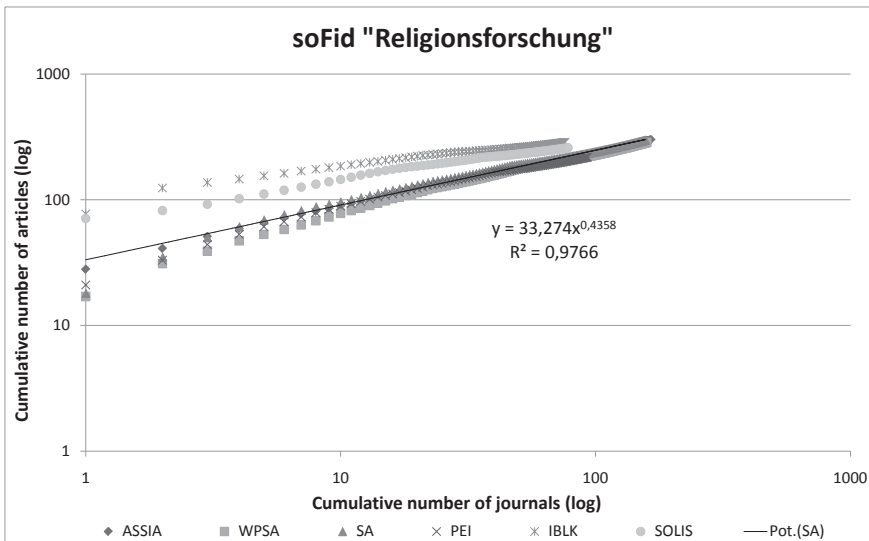


Abbildung 6-5: Die Verteilung der Zeitschriftenartikel für die Deskriptoren-Anfrage des soFid-Themas „Religionsforschung“ in den Datenbanken (doppelt-logarithmische Auftragung)

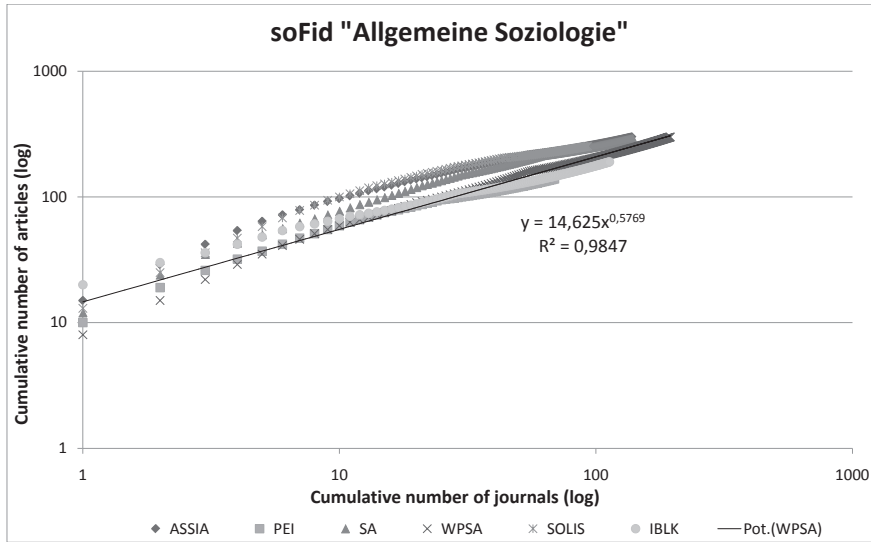


Abbildung 6-6: Die Verteilung der Zeitschriftenartikel für die Deskriptoren-Anfrage des soFid-Themas „Allgemeine Soziologie“ in den Datenbanken (doppelt-logarithmische Auftrugung)

Die Analyse verdeutlicht, dass breit operationalisierte soFid-Themen (Anfrage mit mehreren logisch OR-verknüpften Deskriptoren) ebenfalls die typischen Bradfordschen Verteilungen beim Dokumenttyp Zeitschriftenartikel generieren.

Die Ergebnisse der empirischen Tests zu Ein-Wort-Anfragen (Abschnitt 6.2) bzw. zu den Deskriptoren-Anfragen der soFid-Themen (Abschnitt 6.1) verdeutlichen, dass Bradford-Verteilungen sehr konstant für unterschiedlichste Recherchen in Fachdatenbanken nachzuweisen sind (vgl. dazu These 2 in Kapitel 10).

6.4 Analyse der „Informetrics“-Literatur in der LISA-Datenbank

Die folgende Studie beschreibt eine weitere Anwendung von Bradfordizing in einer Fachdatenbank am Beispiel des Forschungsgebiets „Informetrics“ (dt. Informatie). Die Untersuchung, die ebenfalls auf einer retrospektiven Datenbankrecherche basiert, wurde im Jahr 2006 durchgeführt. Die Ergebnisse der Studie wurden bereits veröffentlicht (siehe (Mayr und Umstätter, 2007, Mayr und Umstätter, 2008)).

Die Studie ist für diese Arbeit relevant, weil die Ergebnisse zum einen erneut die Gültigkeit der Bradfordschen Gesetzmäßigkeit für eine fachspezifische Recherche in Fachdatenbanken verdeutlichen (vgl. These 2 in Kapitel 10) und zum anderen, weil ein spezifischer Mehrwert des Verfahrens Bradfordizing, das hier angewendet wurde, sehr anschaulich an einem realen Beispiel demonstriert werden kann (vgl. These 3 in Kapitel 10).

Die Studie zeigt anhand einer Bradfordizing-Analyse die Kernzeitschriften in diesem Feld und bestätigt auf der Basis des analysierten Datensatzes, dass die Gründung einer neuen Zeitschrift, „Journals of Informetrics“ (JoI⁷⁸), im Jahr 2007 vermutlich zur rechten Zeit erfolgte. Leo Egghe (Egghe, 2005), der Herausgeber der Zeitschrift JoI, belegt zuvor mit mehreren Studien (z. B. (Hood und Wilson, 2001)), dass die Zahl an Veröffentlichungen und die ihrer Autoren in der Informetrie angewachsen ist. Nicht zuletzt durch die Multi- bzw. Interdisziplinarität einer jeden Wissenschaft, die sich an Bradford's Law of Scattering gut ablesen lässt (Umstätter, 1999), hat sich dieses Themengebiet ausgeweitet, was z. B. auch durch die neuen Erkenntnisse in den Webometrics und anderen webbasierten informetrischen Methoden erkennbar ist (siehe z. B. den Übersichtsartikel von (Björneborn und Ingwersen, 2001)).

In der Studie sind insgesamt 3.889 Datensätze analysiert worden, die im Zeitraum 1976 - 2004 in der Datenbank Library and Information Science Abstracts (LISA⁷⁹) im Forschungsbereich der „Informetrie“ nachgewiesen sind und das Wachstum auf diesem Gebiet belegen (siehe Abbildung 6-7).

Um aus dem Bereich der Informetrie eine brauchbare Menge von Daten erfassen zu können, wurde in der LISA-Datenbank nach folgenden Begriffen recherchiert (vgl. Operationalisierung der soFid-Themen in Abschnitt 6.3):

Die Suchanfrage im Freitext der Datenbank

bibliomet* OR cybermet* OR infomet* OR informat* OR scientomet* OR webomet*

generierte insgesamt 3.889 Records, die in LISA (1976 - 2004, CD-ROM-Version) indiziert waren. Die Dokumente sind wie folgt zu den Suchtermen verteilt (vgl. Tabelle 29):

Tabelle 29: Verteilung der Dokumente zu den Anfragetermen in LISA

| Anfrageterme | Gesamtzahl an Dokumenten |
|--------------|--------------------------|
| bibliomet* | 2.851 |
| scientomet* | 1.631 |
| informat* | 292 |
| webomet* | 18 |
| infomet* | 15 |
| cybermet* | 7 |

Von den 3.889 LISA-Records waren 2.414 Dokumente mit „Bibliometrics“, 629 Dokumente mit „Scientometrics“, 90 Dokumente mit „Informetrics“ und 11 Dokumente mit „Webometrics“ indiziert (LISA-Deskriptoren).

78 <http://www.elsevier.com/locate/joi>

79 <http://www.csa.com/factsheets/lisa-set-c.php>

Um die Kernzeitschriften dieses Fachgebietes zu ermitteln, wurden die Publikationen nach der Häufigkeit, in der sie in den verschiedenen Zeitschriften vorkamen, sortiert. Eine typische Power Law-Verteilung ist das Ergebnis dieser Analyse (vgl. Abbildung 6-7).

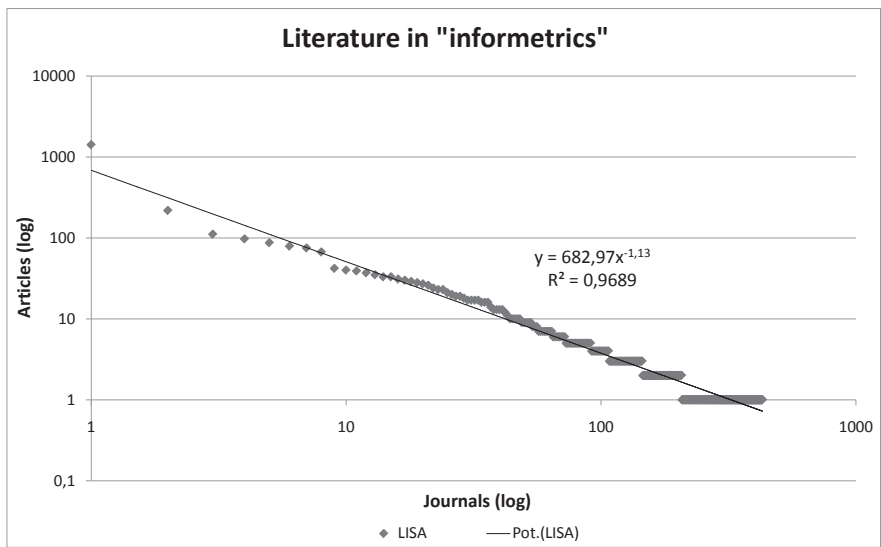


Abbildung 6-7: Power Law-Darstellung der Publikationen in „informetrics“ (Daten aus LISA). Doppelt-logarithmische Skalierung.

Die Kernzeitschriften nach Bradfordizing sind folgende (siehe Tabelle 30):

Tabelle 30: Kernzeitschriften zum Thema „informetrics“ (Daten aus LISA)

| Zeitschrift | Anzahl der Artikel |
|---|--------------------|
| Scientometrics | 1.413 |
| Journal of the American Society for Information Science | 218 |
| Nauchno Tekhnicheskaya Informatsiya | 110 |
| Revista Espanola de Documentacion Cientifica | 96 |
| Journal of Information Science | 87 |
| Information Processing and Management | 79 |
| Journal of Documentation | 75 |
| Annals of Library Science and Documentation | 66 |

Die Bradfordizing-Analyse zeigt (siehe Abbildung 6-7 und Tabelle 30), dass die Zahl der Aufsätze in der Zeitschrift „Scientometrics“ etwa 2- bis 4-mal höher liegt, als es aus dem

weiteren Verlauf der Bradford-Verteilung zu erwarten wäre. Dieses Ergebnis kann als ein Indiz dafür gewertet werden, dass möglicherweise Raum für eine weitere Kernzeitschrift im Bereich der Informetrie ist.

Abbildung 6-8 visualisiert die Daten in der für Bradford charakteristischen Weise. Die y-Achse enthält die kumulierte Anzahl der Artikel der einzelnen Zeitschriften. Die Daten sind einmal mit der häufigsten Zeitschrift „Scientometrics“ und einmal ohne diese Zeitschrift aufgetragen.

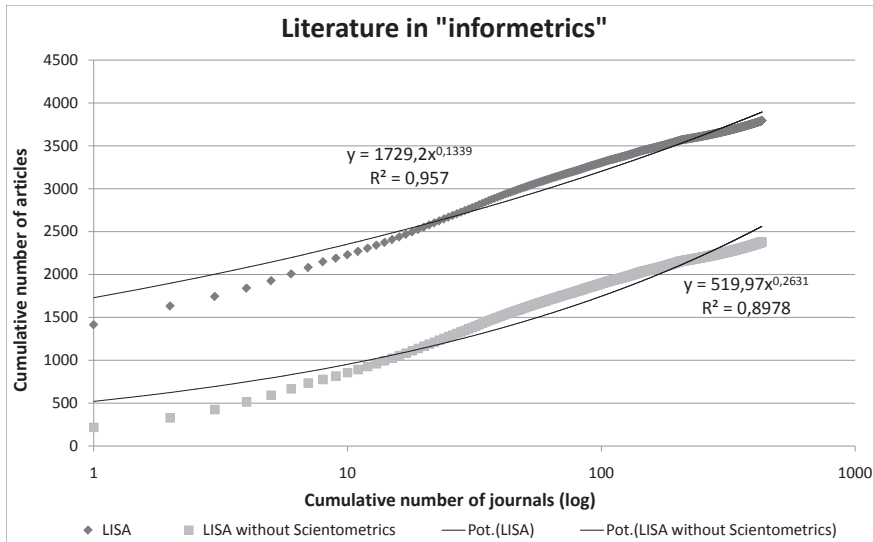


Abbildung 6-8: Bradford-Verteilung der Publikationen in „informetrics“ (Daten aus LISA). Halblogarithmische Darstellung.

Beginnend mit dem Klassiker „Statistical Bibliography“ von E. Windham Hulme (1923) als dem Pionier auf dem Gebiet und nachfolgend mit Bradford (1934), Lotka, Otlet, Pritchard, Zipf und weiteren Autoren, erscheint die Verdopplungsrate der Informetrics bei LISA mit etwa 10 Jahren (siehe Abbildung 6-9). **Eine grobe Abschätzung aller Publikationen im Fachgebiet der „informetrics“** ergibt damit eine Verdopplungsrate, die etwa zweimal so hoch ist wie die der wissenschaftlichen Literatur insgesamt. Da LISA (Rautensymbol in Abbildung 6-9), wie vergleichbare Datenbanken, nur schätzungsweise ~ 40 % des Gesamtaufkommens der Literatur dieses Fachgebietes erfasst, dürften die realen Zahlen im Jahre 2002 bei etwa 500 Publikationen gelegen haben (siehe Quadrate in Abbildung 6-9). Die ursprünglichen LISA-Werte wurden dazu mit 2,5 multipliziert.

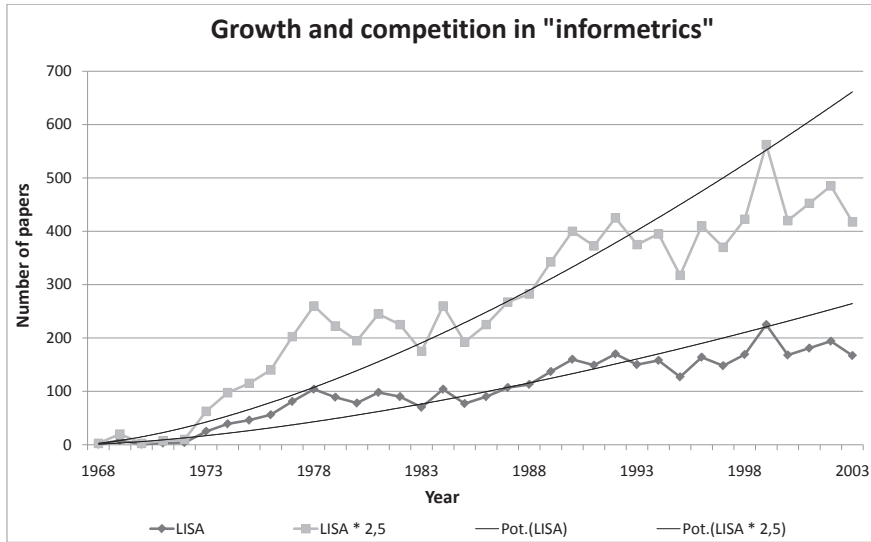


Abbildung 6-9: Wachstum der Publikationen in der Informetrie. Raute: Nachweise in LISA (Daten aus LISA). Quadrate: Schätzung der realen Zahl (Nachweise in LISA * Faktor 2,5)

Die hier skizzierte Vorstudie demonstriert die Potenziale und den Nutzen der Bradfordizing-Analyse (vgl. These 3 in Kapitel 10). Neben der Bestätigung der Gesetzmäßigkeit für eine retrospektive Recherche in einer Fachdatenbank werden weitere Mehrwerte dieses bibliometrischen Verfahrens sichtbar. Ein Nutzer, der nur wenige prominente Begriffe eines Forschungsgebiets kennt, erhält über die Bradfordizing-Analyse Zugriff auf die Kernzeitschriften dieses Gebiets (vgl. Tabelle 30). Des Weiteren kann auf Basis des „zu hohen“ Werts der Zeitschrift „Scientometrics“ im LISA-Datensatz die These aufgestellt werden, dass die Gründung einer neuen Zeitschrift für das untersuchte Gebiet wahrscheinlich angezeigt erscheint.

Die Ergebnisse der bibliometrischen Untersuchung der LISA-Datenbank stützen damit mehrere Thesen dieser Arbeit.

Die folgende Untersuchung der Periodizität wurde im Anschluss an die Relevanzauszählungen (siehe Abschnitt 5.7) durchgeführt und hat hauptsächlich explorativen Charakter.

6.5 Periodizität der Zeitschriften

Die Analyse der Periodizität oder Erscheinungsweise der analysierten Zeitschriften in den Tests stellt eine Anschlussuntersuchung dar, die unabhängig von den zuvor dargestellten Relevanzbewertungen zu sehen ist. Ziel der Untersuchung ist es, herauszufinden, wie häufig die einzelnen Zeitschriften über alle 25 Topics eines CLEF-Jahrgangs in den

Dokumentenpools vorkommen. Die Frage ist, inwiefern sich die Zeitschriften in den drei Zonen (core, z2 und z3) bzgl. ihrer Erscheinungsweise unterscheiden.

Um die Erscheinungsweise der Zeitschriften pro Zone zu untersuchen und den Auswertungsaufwand möglichst gering zu halten, wurden alle analysierten Zeitschriftenartikel zu den 25 Topics eines CLEF-Jahrgangs zusammengespielt. Dieses Verfahren wurde für die CLEF-Jahrgänge 2003, 2005 und 2007 angewendet. Die Zeitschriftenartikel eines kompletten Jahrgangs wurden anschließend nach dem Verfahren Bradfordizing umorganisiert und in drei Zonen (core, z2 und z3) eingeteilt. Da die Zeitschriften, die in der Datenbank SOLIS ausgewertet werden, i. d. R. einen Vermerk zur Erscheinungsweise tragen, konnte die mittlere Periodizität bzw. Erscheinungsweise der Zeitschriften in einer Zone bestimmt werden. Tabelle 31 zeigt ein Beispiel für Zeitschriften, die über ihre ISSN-Nummer identifiziert werden, und die Periodizität dieser Zeitschriften. Die Periodizität bezieht sich darauf, wie häufig eine Zeitschrift pro Jahr erscheint. In SOLIS finden sich auch Zeitschriften, für die keine Erscheinungsweise vermerkt ist. Diese Zeitschriften wurden bei der Bestimmung der Periodizität nicht berücksichtigt (siehe auch Beispiel in Tabelle 31).

Tabelle 31: ISSN-Nummer (Zeitschrift) und deren Periodizität

| ISSN | Periodizität bzw. Erscheinungsweise pro Jahr | |
|-----------|---|---------------|
| 0341-4515 | 6 | zweimonatlich |
| 0170-1754 | 12 | monatlich |
| 0723-8525 | 52 | wöchentlich |
| 0171-3957 | 4 | quartalsweise |
| 0941-5491 | k. A. | |

Die Erscheinungsweise der Zeitschriften je Zone wurde aus Zeitgründen nur für die CLEF-Jahrgänge 2003, 2005 und 2007 berechnet und miteinander verglichen.

Erscheinen Kernzeitschriften (rein quantitative Analyse) häufiger als Zeitschriften in den nachfolgenden Zonen?

Die Tabelle 32 präsentiert die Auswertungen der Erscheinungsweise der Zeitschriften der drei CLEF-Evaluationsreihen der Jahre 2003, 2005 und 2007. Es wurde jeweils die gesamte Zeitschriftenliteratur der 25 Topics eines gesamten Jahrgangs zusammen ausgewertet.

Die Auswertungen zeigen, dass die Zeitschriften im Core im Durchschnitt (Makro-Durchschnitt) monatlich (ca. 12-mal im Jahr) erscheinen, während die Zeitschriften in den nachfolgenden Zonen im Durchschnitt alle zwei Monate (ca. 7-mal im Jahr) erscheinen. Die Werte zwischen den einzelnen Jahrgängen unterscheiden sich in nicht nennenswertem Umfang.

Tabelle 32: Durchschnittliche Erscheinungsweise der Zeitschriften der Datenbank SOLIS für die 25 Fragestellungen der drei CLEF-Jahrgänge. Durchschnittliche Zahl der Artikel pro Zeitschrift im Core, z2 und z3.

| Jahr | Zeitschriften | Erscheinungsweise pro Jahr | Artikel pro Heft 10 Jahre | Artikel pro Heft |
|------|---------------|-------------------------------|------------------------------|------------------|
| 2003 | core | 12,4 | 8,8 | 0,88 |
| 2003 | z2 | 7,1 | 3,8 | 0,38 |
| 2003 | z3 | 6,7 | 0,8 | 0,08 |
| 2005 | core | 12,3 | 11,6 | 1,16 |
| 2005 | z2 | 7,5 | 4,2 | 0,42 |
| 2005 | z3 | 7,3 | 0,7 | 0,07 |
| 2007 | core | 14,2 | 8,1 | 0,81 |
| 2007 | z2 | 7,6 | 3,4 | 0,34 |
| 2007 | z3 | 6,9 | 0,7 | 0,07 |

6.6 Informetrische Analyse von Top Level Domains

Die folgende Studie ist inhaltlich am weitesten vom Thema dieser Arbeit entfernt und soll lediglich Hinweise darauf geben, ob sich die Gesetzmäßigkeiten, die sich in bibliographischen Datenbanken beobachten lassen (vgl. Abschnitte 6.2 - 6.4), auf die Situation im World Wide Web übertragen lassen. Dies soll am Beispiel der Verteilung von Top Level Domains im Internet demonstriert werden.

Die Analyse von Daten, die über Anfragen bei Suchmaschinen generiert werden, hat sich für diese Voruntersuchung angeboten. Exemplarisch sollen hier zunächst zwei Beispielanwendungen genannt werden, die Suchmaschinen-Daten verarbeiten und ebenfalls informatrische bzw. webometrische Analysen mit diesen Daten durchführen:

- Webometrics Ranking of World's Universities⁸⁰: Das Webometrics Ranking listet Universitäts-Websites und Dokumenten-Repositories auf Basis einer Reihe aus dem Web generierter Angaben. In dem Ranking wird unter anderem die Anzahl der erfassten Webseiten durch die Suchmaschine Google verarbeitet.
- Publish or Perish (PoP⁸¹): PoP ist eine Client-Software, die Google Scholar-Ergebnisdaten parst und auf Basis der Zitationsdaten der Treffer Indikatoren berechnet.

Die Suchmaschinen-Schnittstellen (APIs) der kommerziellen Suchmaschinenanbieter sind ein Instrument, um informatrische Untersuchungen mit Internet-Dokumenten durchzuführen und sollen hier exemplarisch für Analysen mit Suchmaschinen-Daten dargestellt

80 <http://www.webometrics.info/>
81 <http://www.harzing.com/resources.htm#/pop.htm>

werden (vgl. (Bar-Ilan, 1998/9, Stein, 2002, Tosques und Mayr, 2005, Tosques und Mayr, 2008)). Die Analyse von Trefferergebnissen kommerzieller Internet-Suchmaschinen gilt seit einigen Jahren als ein Instrument der Webometrie. Verschiedene webometrische Untersuchungen haben Aufmerksamkeit in der Fachwelt hervorgerufen (z. B. (Rousseau, 1997, Thelwall et al., 2005)).

Ziel dieses Abschnittes ist es, anhand von einfachen Tests die Einsatzmöglichkeiten, aber auch Einschränkungen der Suchmaschinen-Schnittstellen (am Beispiel der Google SOAP API⁸²) aufzuzeigen.

Die Domainname-Analyse der Top Level Domains (TLD) aus den Treffer-URLs ist eine Implementation einer typischen webometrischen Analyse, die sehr ähnlich wie die Bradfordizing-Analyse funktioniert. Die Beispielsanwendung⁸³ zeigt die Verteilung der TLDs innerhalb der zurückgelieferten Treffermenge (siehe Tabelle 33). Die Analyse der Top Level Domain-Häufigkeiten in Suchmaschinen-Trefferlisten lehnt sich thematisch an die Untersuchung von Rousseau an (1997). Rousseau konnte 1997 in der Verteilung der TLDs erstmals die Lotka-Funktion bzw. eine Power Law-Verteilung nachweisen.

Die Analyse der Top Level Domains (TLD) zeigt, dass die Daten der Anfrage nach Lotka verteilt ist bzw. dem Lotka Law folgen (siehe Tabelle 33). Lotka Law und Bradford Law sind sich sehr ähnlich.

Tabelle 33: Verteilung der Top Level Domains (TLD) für die Anfrage `bibliometrics OR informetrics OR scientometrics` über die alternative Suchmaschinen-Schnittstelle Google Web APIs

| Anzahl der TLDs | TLDs | Häufigkeit der TLDs |
|-----------------|---|---------------------|
| 1 | edu | 36 |
| 1 | com | 27 |
| 1 | org | 17 |
| 1 | uk | 14 |
| 1 | de | 9 |
| 1 | ca | 8 |
| 1 | net | 7 |
| 3 | be, fi, nl | 5 |
| 1 | au | 4 |
| 4 | cn, fr, in, se | 3 |
| 9 | br, ch, dk, es, hu, il, info, it, ro | 2 |
| 10 | bg, cl, co, gov, ir, jp, li, no, pl, us | 1 |

82 <http://code.google.com/apis/soapsearch/reference.html>

83 http://141.20.126.11/cgi-bin-gs/tld_chart.cgi

Der folgende Screenshot (siehe Abbildung 6-10) zeigt, wie die Auswertung der TLDs on-the-fly realisiert werden kann und die Daten im Anschluss graphisch und tabellarisch verarbeitet werden können.

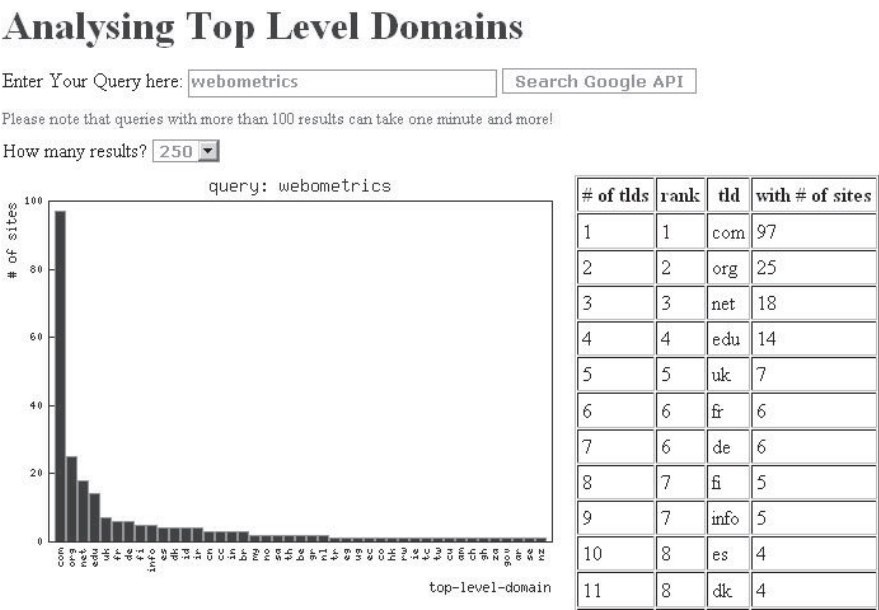


Abbildung 6-10: Analyse der Anfrage „webometrics“ nach Top Level Domains

Suchmaschinen-APIs werden zunehmend zur Datenbeschaffung (Data gathering) für wissenschaftliche Studien verwendet. Die Google APIs sind beispielsweise in den letzten Jahren in Forschungsprojekten der Arbeitsgruppen von Thomas Krichel, Mike Thelwall (Thelwall, 2007), Stevan Harnad (Harnad et al., 2004), Frank McCown (McCown und Nelson, 2007) oder Steffen Staab (Cimiano et al., 2003) eingesetzt worden.

Informetrische Gesetzmäßigkeiten konnten zumindest in Stichproben über die Suchmaschinen-Ergebnisse bestätigt werden (Mayr und Tosques, 2005a, Mayr und Tosques, 2005b). Insbesondere die Verteilung der TLDs in Form der typischen Long-Tail-Distributions (Power Laws) konnte in den Treffern der Google API nachgewiesen werden. Damit lässt sich das Verfahren Bradfordizing prinzipiell auch auf Internet-Dokumente anwenden.

Einschränkend muss bemerkt werden, dass Transparenz und Nachvollziehbarkeit der Ergebnisausgabe bei den kommerziellen Suchmaschinen nicht immer gegeben sind. Außerdem bringen die spezifischen Gegebenheiten im Web bzw. bei der Analyse von Suchmaschinen-Daten (z. B. relativ wenige verfügbare TLD oder Spam/Werbung in den Daten) weitere Probleme mit sich.

Nachdem in diesem Kapitel typische Effekte der Anwendung von Crosskonkordanzen und Bradfordizing beschrieben wurden, sollen im folgenden Kapitel die Hauptergebnisse dieser beiden Verfahren dargestellt werden.

7 Ergebnisse

Das folgende Kapitel führt die Ergebnisse der Experimente zu den eingangs (siehe Abschnitt 1.3) ausgeführten Forschungsfragestellungen zusammen. Die zugrunde liegenden Forschungsfragen werden jeweils der Ergebnisdarstellung vorangestellt. Die Fragestellungen in diesem Kapitel konkretisieren die vergleichsweise offenen Fragestellungen der Vorstudien im vorherigen Kapitel 6. Querbezüge zwischen den Untersuchungsergebnissen der Vorstudien und der Hauptuntersuchung werden erwähnt. Die Ergebnisdarstellung ist bewusst knapp gehalten, da die Hauptergebnisse in der anschließenden Diskussion (siehe Kapitel 8) wieder aufgegriffen und vertiefend erörtert werden. Die gesammelten Ergebnisthesen dieser Studie finden sich in Kapitel 10.

7.1 Effekte der Crosskonkordanzen beim IR

Welche Auswirkungen und Effekte hat das Konzept semantischer Heterogenitätsbehandlung (z. B. Crosskonkordanzen) für die Suche nach Literatur?

Der in Abschnitt 5.2.1 beschriebene Retrieval-Test des Projekts KoMoHe zeigt, dass eine Suche, die über die Komponenten der Heterogenitätsbehandlung (Crosskonkordanzen) transformiert wird, das Suchergebnis signifikant verbessern kann. Zwei Information Retrieval-Tests wurden festgelegt, um die Qualität der Crosskonkordanzen bei der Suche zu evaluieren.

- Test A: Kann der Einsatz der Term-Mappings (Crosskonkordanzen) die Suche gegenüber einer nichttransformierten fachlichen (z. B. über ein kontrolliertes Vokabular) Suche verbessern?

Test A evaluierte, ob die Ersetzung einer Anfrage aus Termen des Vokabulars A (CT) durch kontrollierte Terme aus dem Vokabular B (Transformationen durch Term-Mappings, siehe TT) die Suche in Datenbank B verbessern kann. Wenn die Term-Transformation unpräzise oder doppeldeutig ist oder sich die Vokabulare überlappen, dann kann die Übersetzung der Originalanfrage in die abgebildete Anfrage „Rauschen“ bzw. Vagheit in die Anfrageformulierung bringen, was sich negativ auf die Qualität der Suche auswirken kann.

- Test B: Kann der Einsatz der Term-Mappings (Crosskonkordanzen) die Freitext-Suche verbessern?

Test B evaluierte, ob das Hinzufügen von kontrollierten Termen, die aus dem Mapping natürlichsprachiger Anfrageterme zu dem kontrollierten Vokabular einer Datenbank (FT-CK) gewonnen wurden, zu einer Freitext-Anfrage (FT) die Retrieval-Ergebnisse verbessern kann. Einige der individuellen Anfragen in den Tests konnten nicht erweitert werden, weil keine passenden kontrollierten Terme gefunden werden konnten.

Nachfolgend finden sich die Evaluationsergebnisse⁸⁴ der beiden Test-Reihen.

84 Dieser Abschnitt basiert auf dem Artikel Mayr, Philipp und Petras, Vivien (2008): Cross-concordances: terminology mapping and its effectiveness for information retrieval, IFLA World

7.1.1 Test A: Suche mit kontrollierten Termen

Tabelle 34 zeigt einen Überblick über die gemittelten Ergebnisse aller 13 getesteten Crosskonkordanzen. Die gemittelten Werte über alle evaluierten Crosskonkordanzen finden sich in der Ergebnisthese dieser Studie (siehe Kapitel 10).

Die letzte Zeile der Tabellen in diesem Abschnitt zeigt jeweils die Unterschiede zwischen den Anfragearten in Prozentpunkten.

Tabelle 34: Ergebnisse der Test A-Evaluation für alle Crosskonkordanzen (N=13). Der Unterschied der Anfragearten (CT und TT) beim Recall liegt bei 91,8 % (fett). Das bedeutet, dass die Anfrageart TT über alle Crosskonkordanzen 91,8 % mehr relevante Dokumente generiert als die Anfrageart CT.

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|---|-----------|----------|---------|--------|-----------|--------|--------|
| CT | 156,5 | 144,8 | 42,0 | 0,3152 | 0,2214 | 0,1987 | 0,1748 |
| TT | 325,4 | 144,8 | 88,2 | 0,6047 | 0,3391 | 0,3052 | 0,2848 |
| Unterschied zwischen CT und TT in % | | | | 91,8% | 53,2% | 53,6% | 62,9% |

Die Suche, die die Termtransformation nutzt, verdoppelt die Anzahl der gefundenen Dokumente, d. h. mehr Dokumente, die die Anfrageterme enthalten, werden gefunden. Der Recall verbessert sich um 91,8 %, während sich die Precision um 53,2 % verbessert. Der Einsatz einer Crosskonkordanz in dieser bestimmten Suche findet nicht nur mehr relevante Dokumente (Recall), sondern ist zudem noch präziser (Precision) als eine Suche ohne Termtransformation.

Diese enorme Verbesserung geht jedoch teilweise auf die Übersetzung zwischen Englisch und Deutsch in den bilingualen Crosskonkordanzen zurück. Monolinguale Term-Mappings dürften teilweise ineffektiv sein, weil die abgebildeten Terme identisch sind, was bei übersetzten Mappings nicht der Fall ist.

Tabelle 35 zeigt die Retrievalergebnisse, wenn die bilingualen Crosskonkordanzen aus dem Test entfernt werden:

Library and Information Congress, Québec, Canada. URL: http://www.ifla.org/IV/ifla74/papers/129-Mayr_Petras-en.pdf.

Tabelle 35: Ergebnisse der Test A-Evaluation für alle monolingualen Crosskonkordanzen (N=12)

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|--|-----------|----------|---------|--------|-----------|--------|--------|
| CT | 169,6 | 141,2 | 45,5 | 0,3415 | 0,2399 | 0,2153 | 0,1894 |
| TT | 320,5 | 141,2 | 87,6 | 0,6113 | 0,3431 | 0,3126 | 0,2877 |
| Unterschied zwischen CT und TT in % | | | | 79,0% | 43,1% | 45,2% | 51,9% |

Aufgrund der Termüberlappung sollten sich die Retrieval-Ergebnisse für die Crosskonkordanzen zwischen zwei Disziplinen (interdisziplinär) bzw. Crosskonkordanzen innerhalb der gleichen Disziplin (intradisziplinär) unterscheiden. Wenn die Testergebnisse nach Disziplinen aufgeteilt werden, können wir signifikante Unterschiede in den Retrieval-Ergebnissen erkennen. Der Recall und die Precision bei den intradisziplinären Crosskonkordanzen steigen ebenfalls, allerdings weniger stark. Eine kleinere oder negative Änderung der Precision ist aber erwartungsgemäß, da normalerweise beim Information Retrieval Precision und Recall invers miteinander verbunden sind (wenn der Recall steigt, sinkt die Precision).

Tabelle 36 zeigt die durchschnittlichen Precision- und Recall-Werte für alle intradisziplinären Crosskonkordanzen. Für einsprachige intradisziplinäre Crosskonkordanzen steigen Precision und Recall immer noch an, aber sehr viel weniger im Vergleich zur Gesamtheit der Crosskonkordanzen.

Tabelle 36: Ergebnisse der Test A-Evaluation für intradisziplinäre Crosskonkordanzen (N=5)

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|--|-----------|----------|---------|--------|-----------|--------|--------|
| CT | 126,6 | 101,3 | 36,2 | 0,3726 | 0,2491 | 0,2002 | 0,1637 |
| TT | 238,9 | 101,3 | 59,9 | 0,5189 | 0,3335 | 0,2784 | 0,2352 |
| Unterschied zwischen CT und TT in % | | | | 39,3% | 33,9% | 39,1% | 43,7% |

Eine außergewöhnliche Verbesserung von Recall und Precision kann bei der Anwendung der interdisziplinären Crosskonkordanzen beobachtet werden. Recall und Precision steigen deutlicher an als der Durchschnitt der Crosskonkordanzen (siehe Tabelle 37):

Tabelle 37: Ergebnisse der Test A-Evaluation für interdisziplinäre Crosskonkordanzen (N=8)

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|--|-----------|----------|---------|--------|-----------|--------|--------|
| CT | 175,2 | 171,9 | 45,6 | 0,2794 | 0,2041 | 0,1978 | 0,1817 |
| TT | 379,4 | 171,9 | 105,9 | 0,6583 | 0,3426 | 0,3220 | 0,3157 |
| Unterschied zwischen CT und TT in % | | | | 135,6% | 67,8% | 62,8% | 73,7% |

Der Einsatz der Crosskonkordanzen hat mehr als einen positiven Effekt auf die Suche mit kontrollierten Termen. Die Ergebnismenge ist nicht nur größer, sondern auch präziser. Der größte Effekt kann für die Crosskonkordanzen beobachtet werden, die mehr als eine Disziplin umfassen (interdisziplinäre Suche).

7.1.2 Test B: Freitext-Suche

Tabelle 38 zeigt die Retrieval-Ergebnisse für alle 8 getesteten Crosskonkordanzen:

Tabelle 38: Ergebnisse der Test B-Evaluation für alle Crosskonkordanzen (N=8)

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|---|-----------|----------|---------|--------|-----------|--------|--------|
| FT | 155,3 | 106,4 | 56,2 | 0,6026 | 0,4551 | 0,4101 | 0,3682 |
| FT-CK | 266,8 | 106,4 | 72,8 | 0,7273 | 0,3934 | 0,3203 | 0,3083 |
| Unterschied zwischen FT und FT+TT in % | | | | 20,7 | -13,6 | -21,9% | -16,3% |

Die Ergebnisse zeigen, dass nicht nur mehr, sondern auch mehr relevante Dokumente gefunden wurden. Der durchschnittliche Recall erhöht sich um 20 %. Kontrollierte Terme, die zu einer Freitextanfrage hinzugefügt werden, können ganz generell die Retrieval-Ergebnisse verbessern. Es wurde jedoch ein Abfall der Precision beobachtet, der trotzdem nicht so groß ist wie der Anstieg des Recall.

Tabelle 39 zeigt die Retrieval-Ergebnisse für Crosskonkordanzen, die Terme in der gleichen Disziplin abbilden, während Tabelle 40 die Ergebnisse von 2 interdisziplinären Crosskonkordanzen zeigt:

Tabelle 39: Ergebnisse der Test B-Evaluation für intradisziplinäre Crosskonkordanzen (N=6)

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|---|-----------|----------|---------|--------|-----------|--------|--------|
| FT | 163,8 | 115,8 | 60,2 | 0,5934 | 0,5025 | 0,4635 | 0,4090 |
| FT+ TT | 244,9 | 115,8 | 77,1 | 0,7096 | 0,4449 | 0,3826 | 0,3681 |
| Unterschied zwischen FT und FT+TT in % | | | | 19,6 | -11,5 | -17,5% | -10,0% |

Tabelle 40: Ergebnisse der Test B-Evaluation für interdisziplinäre Crosskonkordanzen (N=2)

| | Retrieved | Relevant | Rel_ret | Recall | Precision | P10 | P20 |
|---|-----------|----------|---------|--------|-----------|--------|--------|
| FT | 129,9 | 78,2 | 44,3 | 0,6303 | 0,3129 | 0,2500 | 0,2459 |
| FT+TT | 332,4 | 78,2 | 59,8 | 0,7805 | 0,2388 | 0,1333 | 0,1292 |
| Unterschied zwischen FT und FT+TT in % | | | | 23,8% | -23,7% | -46,7% | -47,5% |

Im Gegensatz zu der Analyse in Test A sind die Unterschiede zwischen intradisziplinären und interdisziplinären Crosskonkordanzen nicht so groß.

Bei beiden Sets erhöht sich der Recall gegenüber der einfachen Freitext-Suche, während die Precision absinkt. Der Recall erhöht sich leicht für die interdisziplinären Crosskonkordanzen, während die Precision stärker abfällt. Dies kann aber auch daran liegen, dass nur 2 interdisziplinäre Crosskonkordanzen evaluiert wurden, was die Ergebnisse verzerren und sicher nicht als Trend gewertet werden kann.

Die Ergebnisse der Experimente mit der Freitext-Suche könnten wahrscheinlich sehr verbessert werden, wenn die kontrollierten Terme und die natürlichsprachigen Terme besser bei der Anfrageformulierung integriert werden würden, anstatt sie nur aneinanderzufügen. Eine Möglichkeit wäre, die natürlichsprachigen Terme (automatisch) a priori in kontrollierte Terme zu übersetzen, damit eine größere Chance besteht, unterschiedliche kontrollierte Vokabulare und Datenbanken aufeinander abzubilden (siehe Ansatz in (Mayr et al., 2008)).

Die ermittelten Recall- und Precision-Werte der beiden Testreihen (Test A und Test B) wurden anschließend Signifikanztests unterzogen. Insgesamt wurden jeweils 82 Topics (siehe Anhang „Signifikanztest KoMoHe IR-Test“) in den beiden Testreihen (Suche mit kontrollierten Termen und Freitextsuche) berücksichtigt und bezüglich Signifikanz überprüft.

7.1.3 Zusammenfassung

Zusammenfassend kann man sagen, dass die Information Retrieval-Experimente einen positiven Effekt des Einsatzes von Crosskonkordanzen bei heterogenen Datenbanken

zeigen. Die Retrieval-Ergebnisse verbessern sich signifikant für alle Crosskonkordanzen; es zeigt sich jedoch, dass interdisziplinäre Crosskonkordanzen einen deutlich stärkeren (positiven) Effekt auf die Suchergebnisse haben. Für alle Crosskonkordanzen in beiden Testszenarios wurden, verglichen mit den Anfragetypen ohne Nutzung der Crosskonkordanzen, mehr relevante Dokumente gefunden; in bestimmten Fällen waren die Suchergebnisse sogar noch präziser (Anstieg der Precision).

Der Einsatz von Crosskonkordanzen und damit die durchschnittliche Erhöhung des Recall für retrospektive Literaturrecherchen kann damit die Reichweite und indirekt die Qualität von informatrischen Untersuchungen verbessern. Die folgende Fragestellung setzt direkt an diesem Punkt an und analysiert typische Dokumentverteilungen, die sich zu spezifischen Themen in unterschiedlichen Fachdatenbanken identifizieren lassen.

7.2 Bradford-Verteilungen in Fachdatenbanken

Inwieweit lassen sich Bradford-Verteilungen in unterschiedlichen Fachgebieten und insbesondere bei Recherchen in Literaturdatenbanken identifizieren?

Lassen sich diese für Zeitschriftenliteratur typischen Verteilungen auch bei anderen Dokumenttypen (z. B. Monographien) beobachten?

Im folgenden Abschnitt wird untersucht, ob sich Bradford-typische Verteilungen zu fachspezifischen Topics (siehe Kapitel 4) beobachten lassen. Ziel der Untersuchung ist es nicht, Bradford's Law of Scattering in den erhobenen Daten möglichst genau zu identifizieren bzw. Abweichungen der Gesetzmäßigkeit näher zu evaluieren (vgl. dazu z. B. (Bonitz, 1980)). Ziel der Untersuchung ist es, eine möglichst umfangreiche Anzahl von standardisierten Fragestellungen in unterschiedlichen Datenbanken und Fachgebieten durch Bradfordizing zu analysieren, um anschließend das Vorkommen der Bradford-typischen Verteilungen verifizieren oder falsifizieren zu können (vgl. Abschnitt 4.1).

Nachfolgend werden die Auswertungsergebnisse der beiden relevanzbewerteten Datensätze (CLEF und KoMoHe) präsentiert. Der Datensatz CLEF umfasst insgesamt 125 Topics des Cross-Language Evaluation Forum (CLEF) der Jahrgänge 2003 bis 2007 und basiert auf sozialwissenschaftlichen Dokumenten des GIRT-Korpus (Datenbank SOLIS und SOFIS) und deren Relevanzbewertungen. Die Topics sind jeweils dem „domain-specific track“ der CLEF-Evaluationsinitiative entnommen.

Der Datensatz KoMoHe umfasst insgesamt 39 Topics der Information Retrieval-Evaluation im Projekt KoMoHe und basiert auf Dokumenten der Datenbanken SOLIS, SoLit, USB Köln OPAC, Psyndex, World Affairs Online, Medline. Die Relevanzbewertungen sind im Jahr 2007 erfolgt.

Die detaillierten Auswertungen der Topics der beiden Datensätze finden sich im Abschnitt 5.4. Eine Übersicht der Zusammensetzung der evaluierten Dokumente (Datenbanken, Dokumenttypen, Disziplinen) sowie die Beschreibung der Topic-Generierung und der Relevanzbewertungen finden sich im Kapitel 5.

7.2.1 Verteilung der Zeitschriftenartikel auf Bradford-Zonen

Wie in Abschnitt 5.1 beschrieben, wurden die Dokumentenpools pro Topic nach dem Verfahren Bradfordizing analysiert und neuorganisiert. Die Auswertungsergebnisse der beiden Datensätze (CLEF und KoMoHe) und der beiden analysierten Dokumenttypen (Zeitschriftenartikel, Monographien) werden nachfolgend jeweils getrennt voneinander dargestellt.

Auswertung der CLEF-Topics

Bei der Aufteilung der Dokumentmenge pro Topic in drei gleich große Zonen (Core, Zone 2 bzw. z2 und Zone 3 bzw. z3) ergibt sich folgende Verteilung der Zeitschriften über alle fünf Jahrgänge des CLEF-Datensatzes bzw. die Topics des KoMoHe-Datensatzes (die Tabellen enthalten jeweils die Mittelwerte der Auszählungen):

Tabelle 41: Durchschnittliche Anzahl der Journals in den Bradford-Zonen (core, z2 und z3) im CLEF-Datensatz (N=125 Topics).

| CLEF-Jahrgang | Topics | Zeitschriften- artikel | Zeitschriften total | core | z2 | z3 | z2/core | z3/z2 |
|---------------|--------|---------------------------|------------------------|------|-------|-------|---------|-------|
| 2003 | 25 | 144 | 64 | 5,20 | 16,52 | 42,68 | 3,18 | 2,58 |
| 2004 | 25 | 115 | 60 | 5,28 | 17,76 | 36,52 | 3,36 | 2,06 |
| 2005 | 25 | 142 | 61 | 4,64 | 17,24 | 39,56 | 3,72 | 2,29 |
| 2006 | 25 | 153 | 75 | 7,08 | 20,76 | 47,40 | 2,93 | 2,28 |
| 2007 | 25 | 170 | 76 | 6,20 | 20,08 | 49,84 | 3,24 | 2,48 |

Im CLEF-Datensatz werden durchschnittlich – über alle 125 Topics gerechnet – rund 6 Zeitschriften benötigt, um die Kernzone auszufüllen (Mittelwert für alle Kernzeitschriften („core“) = 5,68, siehe Tabelle 41). Für die gleiche Dokumentmenge in Zone 2 und Zone 3 werden im Mittel rund 18 Zeitschriften bzw. rund 43 Zeitschriften benötigt. Es befinden sich demnach ca. 3,3-mal mehr Zeitschriften in Zone 2 als im konzentrierten Core. Um die Dokumentmenge in Zone 3 bereitzustellen, werden wiederum ca. 2,3-mal mehr Zeitschriften benötigt als in der vorherigen Zone 2. Diese Verteilung entspricht der typischen Bradford-Verteilung. Im Durchschnitt wurden über den gesamten Zeitraum 145 Zeitschriftenartikel (siehe ‚Zeitschriftenartikel‘ in Tabelle 41) pro Topic analysiert. Diese 145 Zeitschriftenartikel finden sich in durchschnittlich 67 verschiedenen Zeitschriften pro Topic.

Auswertung der KoMoHe-Topics

Die Verteilung der Zeitschriften im Datensatz KoMoHe zeigt ebenfalls den für Bradford-Verteilungen charakteristischen Zuwachs an Zeitschriften in den Zonen 2 und 3. Die 39 Topics lassen sich drei unterschiedlichen Tests zuordnen:

- Test 1 = innerdisziplinärer Test in den Sozialwissenschaften,
- Test 2 = interdisziplinärer Test in den Sozialwissenschaften und verwandten Disziplinen,
- Test 3 = interdisziplinärer Test in Medizin und Psychologie (außerhalb der Sozialwissenschaften).

Durchschnittlich werden über alle 39 Topics des KoMoHe-Datensatzes rund 10 Zeitschriften benötigt, um die Dokumentmenge für die Kernzone zu erreichen. Für die Zone 2 und Zone 3 werden rund 34 Zeitschriften bzw. rund 116 Zeitschriften benötigt. Es befinden sich demnach ca. 3,3-mal mehr Zeitschriften in Zone 2 als im konzentrierten Core. Um die Dokumentmenge in Zone 3 bereitzustellen, werden wiederum ca. 3,42-mal mehr Zeitschriften benötigt als in der vorherigen Zone 2. Diese Verteilung entspricht der typischen Bradford-Verteilung. Im Durchschnitt wurden über den gesamten Zeitraum 382 Zeitschriftenartikel (siehe ‚Zeitschriftenartikel‘ in Tabelle 42) pro Topic analysiert. Diese 382 Zeitschriftenartikel finden sich in durchschnittlich 160 verschiedenen Zeitschriften pro Topic.

Tabelle 42: Durchschnittliche Anzahl der Journals in den Bradford-Zonen (Core, Z2 und Z3) im KoMoHe-Datensatz (N=39 Topics)

| KoMoHe-Reihe | Topics | Zeitschriften- artikel | Zeitschriften total | core | z2 | z3 | z2/core | z3/z2 |
|--------------|--------|---------------------------|------------------------|-------|-------|--------|---------|-------|
| Test 1 | 12 | 449 | 134 | 8,25 | 27,25 | 98,67 | 3,30 | 3,62 |
| Test 2 | 15 | 309 | 96 | 6,47 | 20,47 | 68,67 | 3,16 | 3,36 |
| Test 3 | 12 | 688 | 250 | 15,92 | 54,83 | 179,33 | 3,45 | 3,27 |

Die Verteilung der Zeitschriften für beide Datensätze CLEF und KoMoHe entspricht damit weitestgehend der typischen Bradford-Verteilung. Dies kann durch die folgenden Abbildungen (Abbildung 7-1 bis Abbildung 7-5) illustriert werden.

Im KoMoHe-Datensatz wurden im Mittel mehr als 2,5-mal so viele Zeitschriftenartikel bewertet wie im CLEF-Datensatz (siehe Spalte ‚Zeitschriftenartikel‘).

Die Ergebnisse der Topics in Test 3 fallen etwas aus dem Rahmen. Die Fragestellungen von zwei der 12 Topics generieren deutlich mehr unterschiedliche Zeitschriften als der Durchschnitt der anderen Topics. Die beiden Topics wurden in der Datenbank Medline untersucht.

Die folgenden Abbildungen visualisieren die Bradford-Verteilung der einzelnen Topics. Die Datenpunkte sind in doppelt-logarithmischer Skala aufgetragen. Auf der x-Achse ist die kumulierte Anzahl der Zeitschriften und auf der y-Achse sind die kumulierten Artikelzahlen in diesen Zeitschriften aufgetragen. 25 Topics eines Jahrgangs der CLEF-Evaluation sind jeweils in einer Abbildung zusammen aufgetragen. Die Daten der 39 KoMoHe-Topics sind in gleicher Weise aufgetragen. Die Topics der drei Testgruppen (vgl. Tabelle 42) werden jeweils gemeinsam aufgetragen. Die Topics der drei Testreihen

Test 1 (innerdisziplinärer Test), Test 2 und 3 (zwei interdisziplinäre Tests) sind jeweils in einer Abbildung zusammen aufgetragen.

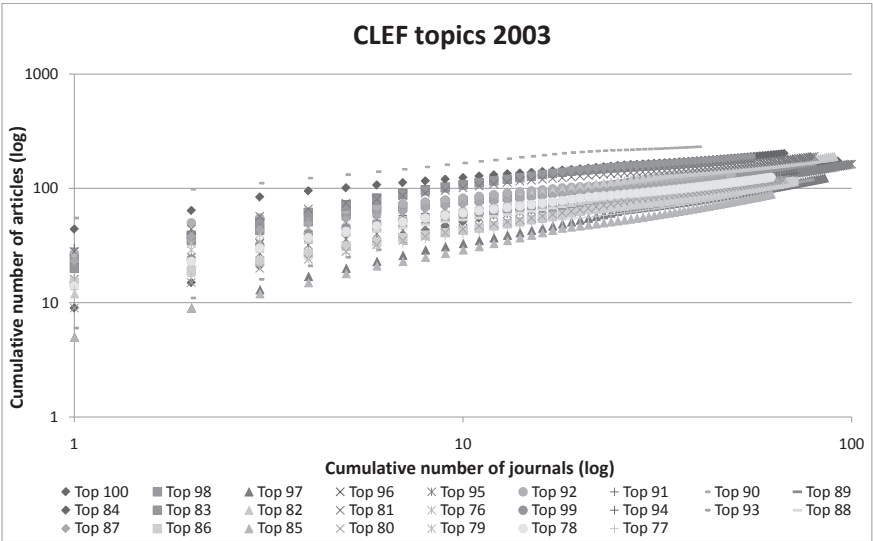


Abbildung 7-1: Bradford-typische Verteilung. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2003. Auf der x-Achse ist die kumulierte Anzahl der Zeitschriften und auf der y-Achse sind die kumulierten Artikelzahlen in diesen Zeitschriften aufgetragen (log-log-Skala).

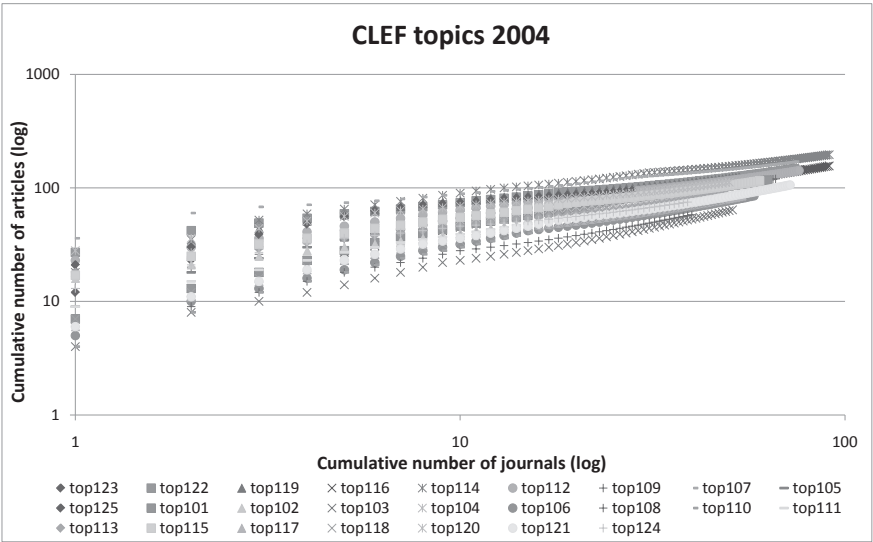


Abbildung 7-2: Bradford-typische Verteilung. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2004 (log-log-Skala).

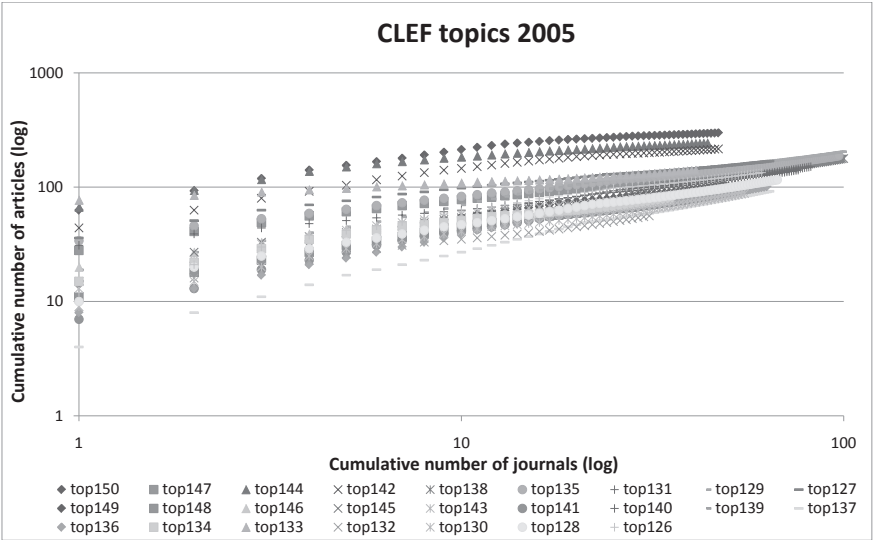


Abbildung 7-3: Bradford-typische Verteilung. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2005 (log-log-Skala).

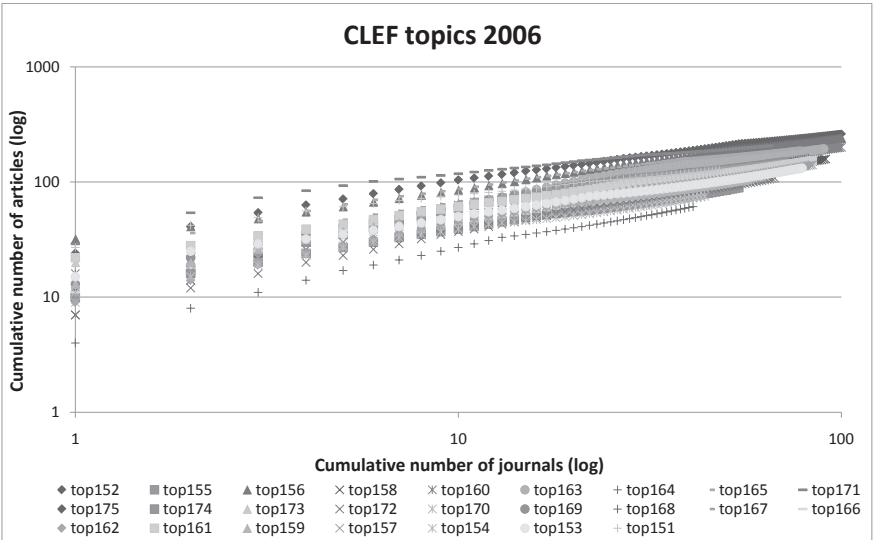


Abbildung 7-4: Bradford-typische Verteilung. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2006 (log-log-Skala).

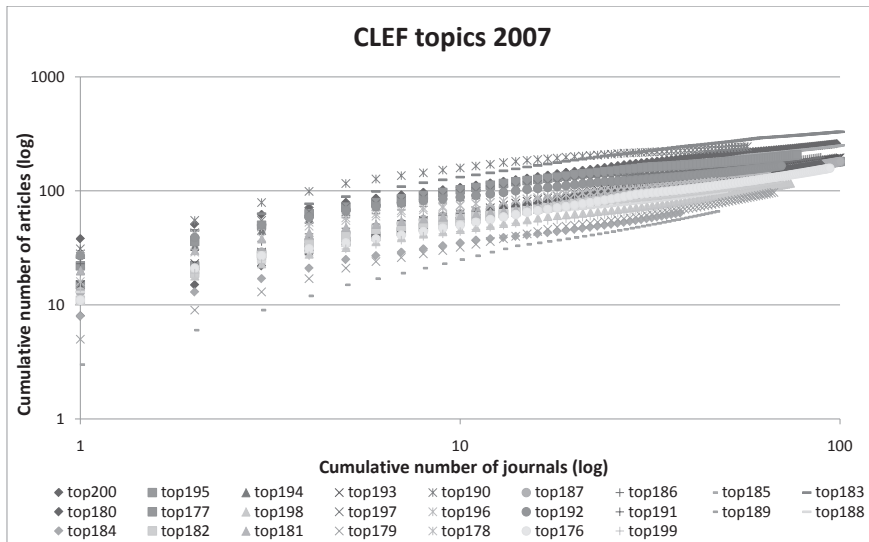


Abbildung 7-5: Bradford-typische Verteilung. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2007 (log-log-Skala).

Die folgenden drei Abbildungen (Abbildung 7-6, Abbildung 7-7 und Abbildung 7-8) visualisieren die Topics der drei KoMoHe-Testreihen.

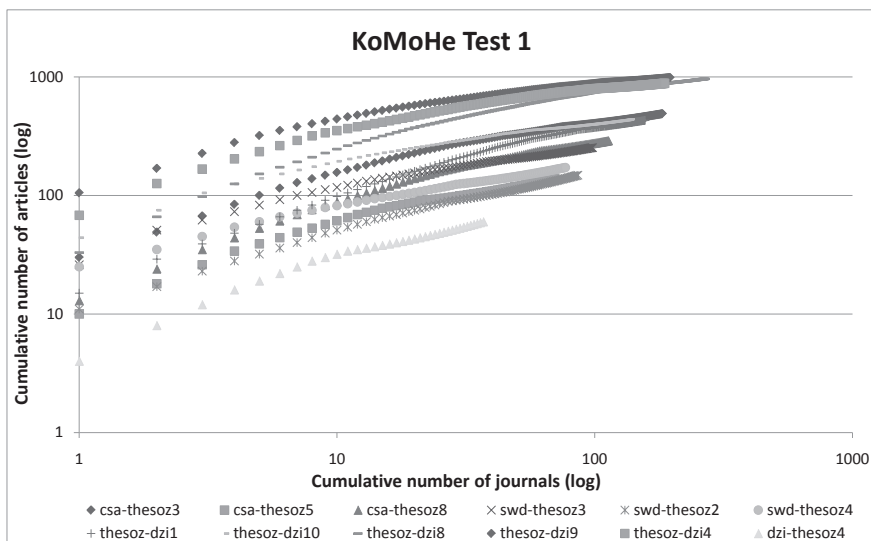


Abbildung 7-6: Bradford-typische Verteilung für 12 Topics der KoMoHe-Evaluation (log-log-Skala). Innerdisziplinärer Test 1 (Zeitschriftenartikel).

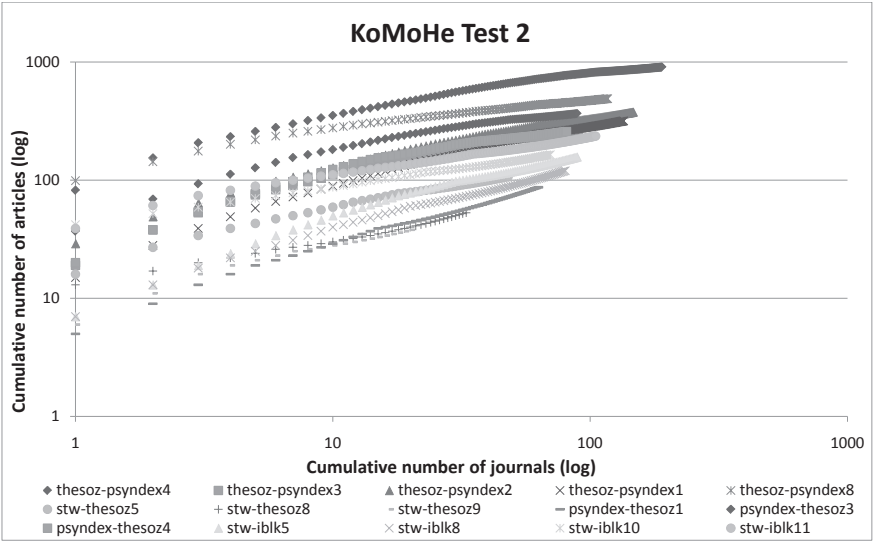


Abbildung 7-7: Bradford-typische Verteilung für 15 Topics der KoMoHe-Evaluation (log-log-Skala). Interdisziplinärer Test 2 (Zeitschriftenartikel).

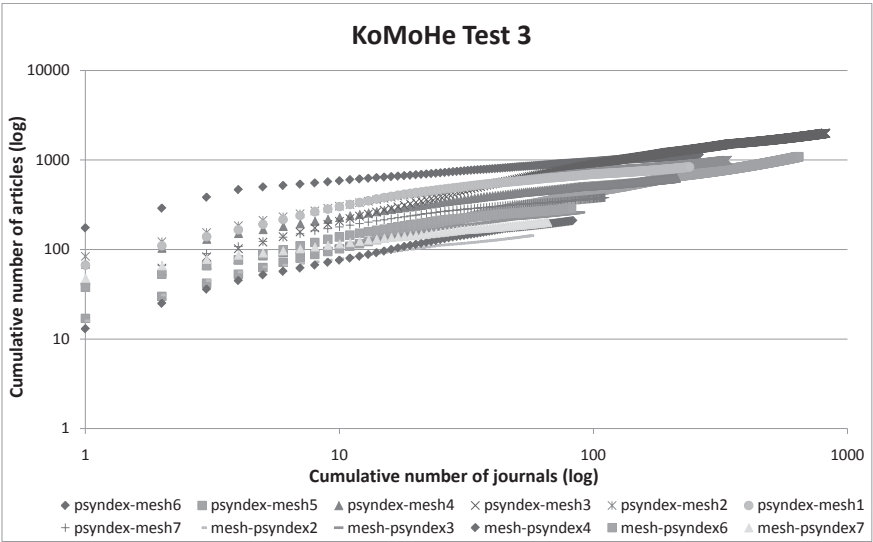


Abbildung 7-8: Bradford-typische Verteilung für 12 Topics der KoMoHe-Evaluation (log-log-Skala). Interdisziplinärer Test 3 (Zeitschriftenartikel).

7.2.2 Verteilung der Monographien auf Bradford-Zonen

Die folgenden beiden Tabellen präsentieren die Vergleichswerte der Analyse der Datensätze CLEF und KoMoHe nach dem Dokumenttyp Monographie. Es wurden ausschließlich Monographien und die Häufigkeit der Verlage, die diese Monographien herausgeben, analysiert.

Auswertung der CLEF-Topics

Im Durchschnitt stellen rund 7 Verlage (Publishers core) die Kernverlage für die analysierten 75 CLEF-Topics dar. Für die Zone 2 werden rund 26 und für die Zone 3 rund 49 Verlage benötigt, um das zweite bzw. dritte Drittel der Monographien eines Topics bereitzustellen⁸⁵. Damit liegen die Werte der Monographien leicht über denen der Zeitschriftenartikel (siehe Tabelle 43).

Tabelle 43: Durchschnittliche Anzahl der Verlage bei Monographien in den Bradford-Zonen (Core, Z2 und Z3) im CLEF-Datensatz (N=75 Topics)

| CLEF | Topics | Monographien | Verlage total | core | z2 | z3 | z2/core | z3/z2 |
|------|--------|--------------|---------------|------|-------|-------|---------|-------|
| 2005 | 25 | 211 | 90 | 8,60 | 29,56 | 51,88 | 3,44 | 1,76 |
| 2006 | 25 | 204 | 77 | 7,08 | 24,24 | 45,76 | 3,42 | 1,89 |
| 2007 | 25 | 235 | 80 | 6,44 | 24,04 | 49,80 | 3,73 | 2,07 |

Auswertung der KoMoHe-Topics

Sehr ähnliche Ergebnisse liefert die Analyse der Monographien und der zugehörigen Verlage im KoMoHe-Datensatz. Im Durchschnitt stellen rund 6 Verlage (Publishers core) die „Kernverlage“ für die analysierten 27 KoMoHe-Topics dar. Für die Zone 2 werden rund 20 und für die Zone 3 rund 48 Verlage benötigt, um das zweite bzw. dritte Drittel der Monographien eines Topics bereitzustellen (siehe Tabelle 44).

85 Hinweis: Analysiert wurden alle mit „book“ ausgezeichneten Dokumente. Darunter befinden sich auch Dokumente, die keine ISBN enthalten. Diese Dokumente wurden der Zone 3 zugeordnet. Damit enthält die Zone 3 im Durchschnitt mehr Dokumente als der Core und Zone 2.

Tabelle 44: Durchschnittliche Anzahl der Verlage bei **Monographien** in den Bradford-Zonen (Core, Z2 und Z3) im KoMoHe-Datensatz (N=27 Topics)

| KoMoHe | Topics | Monographien | Verlage total | core | z2 | z3 | z2/core | z3/z2 |
|--------|--------|--------------|---------------|------|-------|-------|---------|-------|
| Test 1 | 7 | 327 | 97 | 7,29 | 27,14 | 62,57 | 3,73 | 2,31 |
| Test 2 | 15 | 188 | 69 | 5,73 | 19,33 | 44,07 | 3,37 | 2,28 |
| Test 3 | 5 | 119 | 56 | 5,60 | 14,20 | 36,60 | 2,54 | 2,58 |

Die folgenden Abbildungen visualisieren die Bradford-Verteilung der einzelnen Topics für die beiden Evaluationsreihen (CLEF und KoMoHe). Die Datenpunkte sind in doppelt-logarithmischer Skala aufgetragen. Auf der x-Achse ist die kumulierte Anzahl der Verlage und auf der y-Achse sind die kumulierten Monographie-Zahlen dieser Verlage aufgetragen.

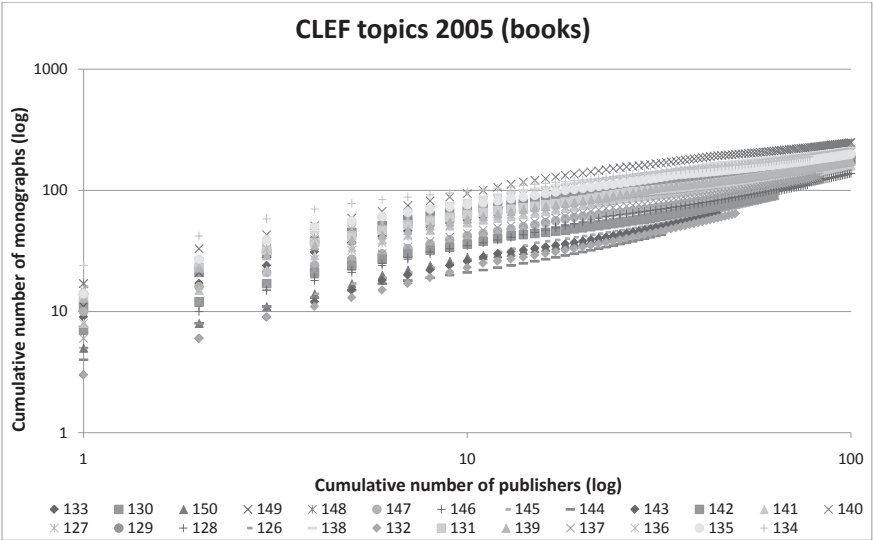


Abbildung 7-9: Bradford-typische Verteilung. Analysiert wurden die **Monographien** der 25 Topics aus dem Jahre 2005. Auf der x-Achse ist die kumulierte Anzahl der Verlage und auf der y-Achse sind die kumulierten Monographie-Zahlen dieser Verlage aufgetragen (log-log-Skala).

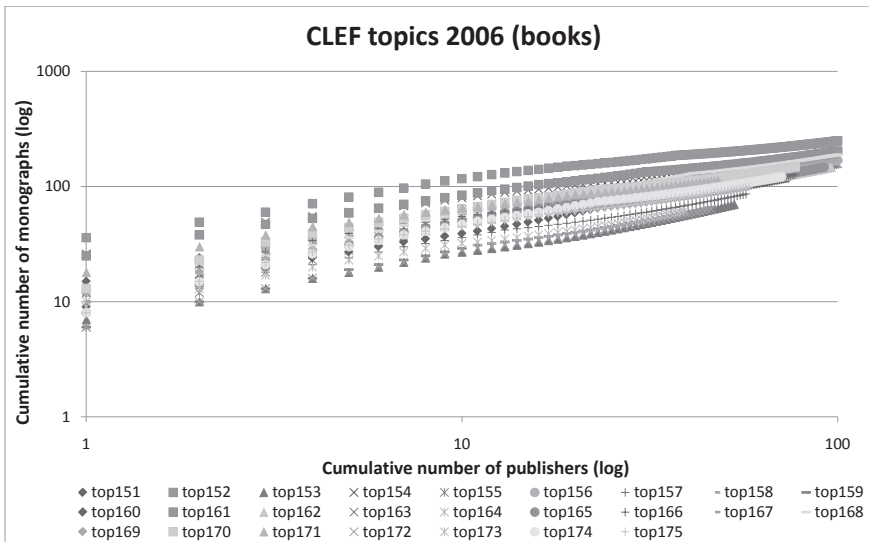


Abbildung 7-10: Bradford-typische Verteilung. Analysiert wurden die Monographien der 25 Topics aus dem Jahre 2006 (log-log-Skala).

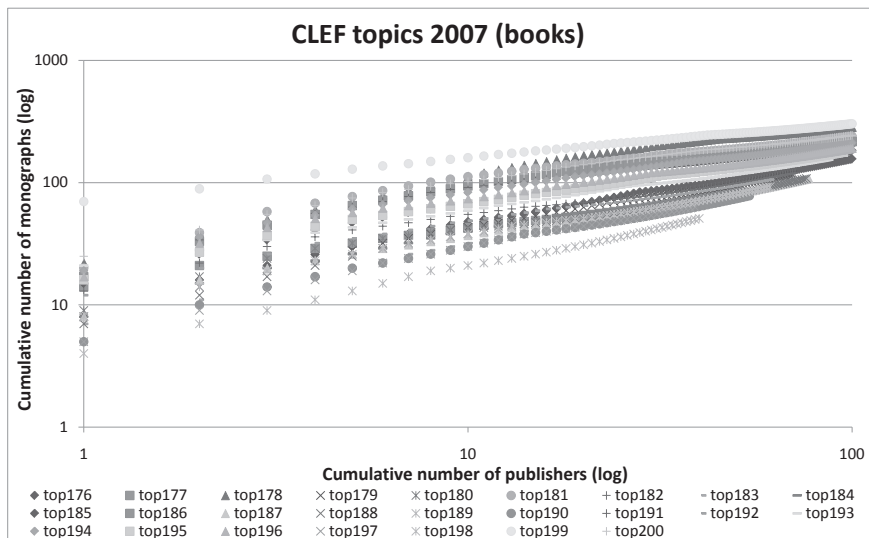


Abbildung 7-11: Bradford-typische Verteilung. Analysiert wurden die Monographien der 25 Topics aus dem Jahre 2007 (log-log-Skala).

Die folgenden drei Abbildungen (Abbildung 7-12, Abbildung 7-13 und Abbildung 7-14) visualisieren die Topics der drei KoMoHe-Testreihen.

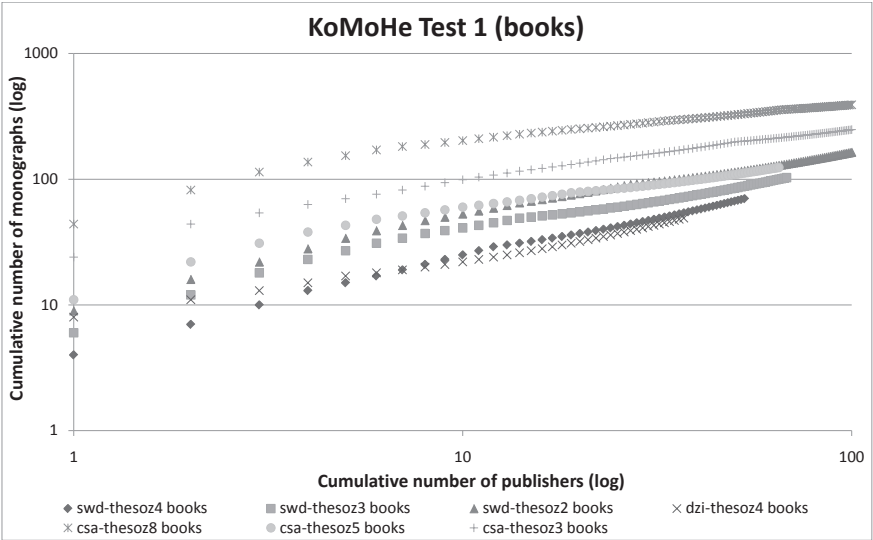


Abbildung 7-12: Bradford-typische Verteilung für 7 Topics der KoMoHe-Evaluation (log-log-Skala). Innerdisziplinärer Test 1 (Monographien).

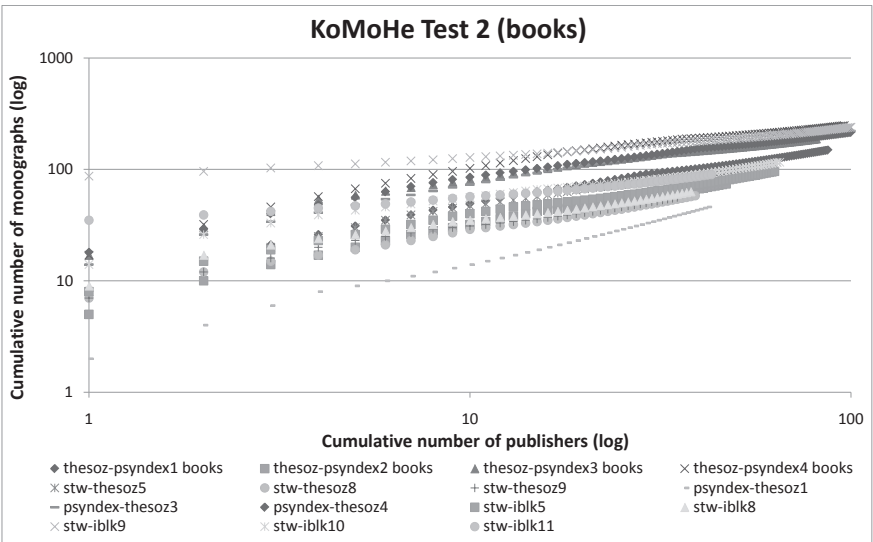


Abbildung 7-13: Bradford-typische Verteilung für 15 Topics der KoMoHe-Evaluation (log-log-Skala). Interdisziplinärer Test 2 (Monographien).

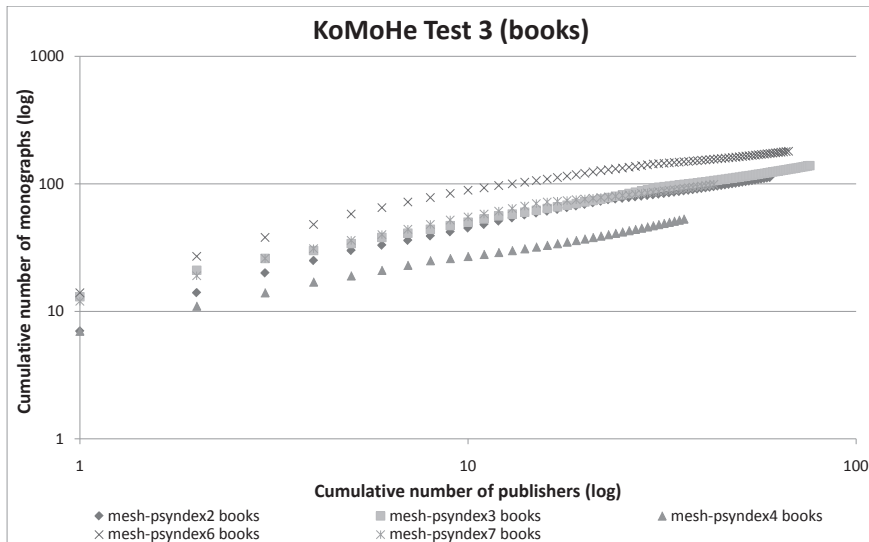


Abbildung 7-14: Bradford-typische Verteilung für 5 Topics der KoMoHe-Evaluation (log-log-Skala). Interdisziplinärer Test 3 (Monographien).

7.2.3 Zusammenfassung

Die Analyse der Topics zeigt deutlich, dass Bradford-ähnliche Verteilungen für beide Datensätze (Datenbanken SOLIS, SoLit, USB Köln OPAC, Psynindex, World Affairs Online, Medline) und beide Dokumenttypen (Zeitschriftenartikel und Monographien) nachgewiesen werden können.

Prinzipiell kann das Verfahren Bradfordizing aufgrund des Auftretens der Bradford-schen Gesetzmäßigkeit für beide Dokumenttypen auf Basis der rein quantitativen Analyse angewendet werden.

Es zeigt sich, dass die Graphen der Topics der CLEF-Evaluationen über die Jahre beim Dokumenttyp Zeitschriftenartikel homogener verlaufen als die Verteilung der Zeitschriftenartikel der KoMoHe-Evaluation.

Bei der Analyse der Steigungen der Bradford-Graphen zeigt sich, dass durchaus Unterschiede zwischen den einzelnen Topics festzustellen sind. Dies soll allerdings nicht als Einschränkung der vorherigen Aussage gewertet werden, sondern diese Varianz liegt im normalen Streubereich. Abbildung 7-15 visualisiert die Verteilung der Exponenten für 100 unterschiedliche Topics unterschiedlicher CLEF-Reihen.

Im Durchschnitt liegt der Exponent (Steigung bzw. power) der Graphen (Topics) in Abbildung 7-15 bei rund 0,49. Es finden sich aber auch Topics mit sehr starker ($>0,65$) und geringer ($<0,35$) Steigung der Graphen. Eine differenzierte Auseinandersetzung der unterschiedlichen Power Laws der analysierten Topics soll in dieser Arbeit nicht durchgeführt werden.

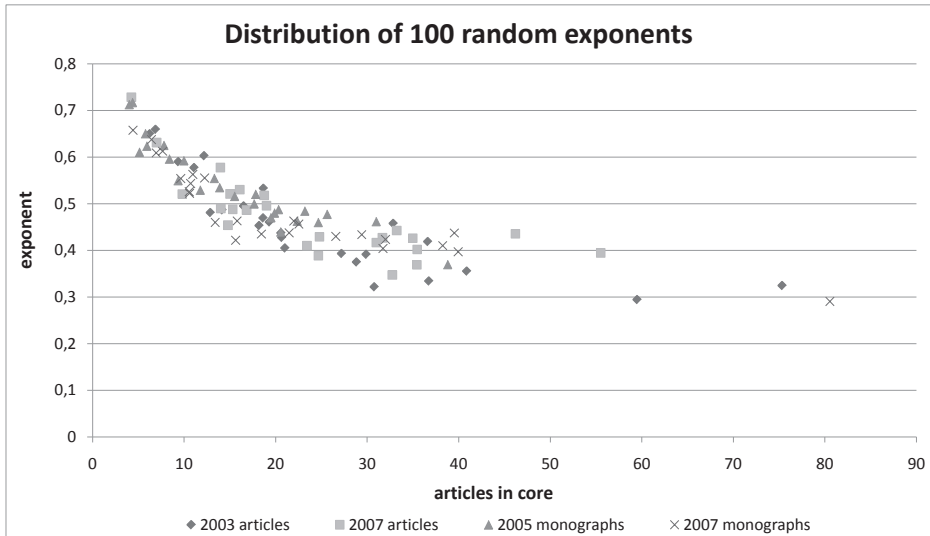


Abbildung 7-15: Abnahme von 100 Exponenten für unterschiedliche Testreihen der CLEF-Evaluation mit zunehmenden „articles in core“

7.3 Befragung zu Bradfordizing als Mehrwertdienst

Lassen sich informatrische Verfahren (z. B. Bradfordizing) als Mehrwertdienste für die Informationssuche in Literaturdatenbanken gewinnbringend nutzen?

Das folgende Interview soll qualitative Hinweise zum Nutzen und der Plausibilität des Verfahrens Bradfordizing beim IR generieren. Es wird angenommen, dass die Hinweise der qualitativen Befragung trotz einer kleinen Stichprobe ausreichen, um neue Interpretationsspielräume für die Diskussion der quantitativen Ergebnisse zu eröffnen. Eine umfangreiche quantitative Analyse dieser Fragestellung schließt sich im nächsten Abschnitt direkt an. Der folgende Abschnitt fasst die Ergebnisse der vier Fragen knapp zusammen.

7.3.1 Interview zu Kernzeitschriften

Von den 25 Topics der Evaluationsrunde 2007 konnten 24 Topics konkreten Personen (N=12) zugeordnet werden (eine Übersicht der Topics findet sich in Tabelle 21). Das Interview besteht aus vier kurzen Fragen, die teils per Telefon, teils in „face to face“-Gesprächen mit den 12 beteiligten Personen gestellt wurden. Ein Beispielfragebogen findet sich im Anhang (Anhang „Beispielfragebogen“).

Nach einer kurzen Einleitung ins Thema und den Zielen des folgenden Interviews wurden vom Autor folgende Fragen gestellt:

Frage 1: Nennen Sie mir bitte spontan mehrere Zeitschriften, die Sie zu den Kernzeitschriften (häufigsten Zeitschriften) des Themas (XY) in der Datenbank SOLIS zählen.

Frage 1 zielt insbesondere auf die Vertrautheit der Personen mit der Zeitschriftenliteratur zu „ihrem“ Thema. Es wurde konkret nach dem Kenntnisstand der Personen gefragt. Die Zeitschriftentitel (Kernzeitschriften), die die Befragten spontan genannt haben, wurden mit protokolliert. Wenn die Personen keine Zeitschriften nennen konnten, wurde dies vermerkt. Die Nennungen wurden während der Befragung mit den gelisteten Kernzeitschriften abgeglichen und direkt an die Befragten rückgekoppelt.

Auf die Frage 1 „spontane Nennung von Kernzeitschriften“ zu einem bekannten Thema konnten zu 13 Topics keine Zeitschriften aufgezählt werden (siehe Tabelle 45). Die Antworten zu den 11 verbleibenden Topics verteilen sich folgendermaßen: Zu 9 Topics konnten eine oder mehrere Kernzeitschriften genannt werden (1 Kernzeitschrift für 4 Topics, 2 Kernzeitschriften für 3 Topics und mehr als 2 Kernzeitschriften für 2 Topics). Zu 8 Topics konnten weitere Zeitschriften genannt werden, die sich allerdings nicht in der Liste der analysierten Kernzeitschriften befanden.

Tabelle 45: Ergebnisse Frage 1. N=24 Topics

| Antworten zu Frage 1 | |
|--|--------|
| „spontane Nennung von Kernzeitschriften“ | Topics |
| keine Nennung | 13 |
| 1 Kernzeitschrift | 4 |
| 2 Kernzeitschriften | 3 |
| mehr als 2 Kernzeitschriften | 2 |
| weitere Zeitschriften | 8 |

Im Anschluss an Frage 1 wurden vom Autor alle durch Bradfordizing bestimmten Kernzeitschriften für das jeweilige Topic vorgelesen⁸⁶.

An das Verlesen der Kernzeitschriften schließt direkt Frage 2 mit einer Bewertung der Kernzeitschriften an.

Frage 2: Hätten Sie diese Zeitschriften als häufigste Veröffentlichungsorgane für Ihr Topic erwartet?

Frage 2 „Erwartungskonformität mit Kernzeitschriften“ fokussiert auf die quantitativ bestimmten Kernzeitschriften zu den einzelnen Topics der Befragung. Frage 2 soll insbesondere klären, ob es sich bei den vorgelesenen Listen der Kernzeitschriften eher um Artefakte der Datenbank handelt, oder ob diese Zeitschriften wirklich als Hauptveröffentlichungsorgane bekannt sind und dementsprechend wahrgenommen werden.

⁸⁶ Im Durchschnitt befanden sich rund 6 Zeitschriften in der Liste der Kernzeitschriften pro Topic (Topics 2007).

Zu 19 Topics wird angegeben, dass diese Zeitschriften zumindest „z. T.“ den Erwartungen der Befragten entsprechen. Von den Interviewpartnern (N=8), die bei Frage 1 keine Kernzeitschrift(en) nennen konnten, wird bei Frage 2 immerhin für 9 Topics mit „ja“ bzw. „z. T.“ geantwortet. Das spricht dafür, dass viele Kernzeitschriften bei den Bearbeitern zwar implizit bekannt sind und als solche wahrgenommen werden, aber in der Situation des Interviews nicht spontan abgerufen wurden. In jedem Fall kann dieses Ergebnis (Nennung der jeweiligen Kernzeitschriften eines Topics) für die befragten Personen als ein Erkenntnisgewinn gewertet werden. Dies wird auch bei den zahlreichen späteren Kommentaren und Hinweisen deutlich. Lediglich zu 5 Topics wurden von den Bearbeitern (N=4) gänzlich andere Kernzeitschriften erwartet (Antwort „nein“) (siehe Tabelle 46).

Tabelle 46: Ergebnisse Frage 2. N=24 Topics

| Antworten zu Frage 2 | |
|---|--------|
| „Erwartungskonformität mit Kernzeitschriften“ | Topics |
| ja | 9 |
| z. T. | 10 |
| nein | 5 |
| weiß nicht | 1 |

Frage 3: Können Sie sich vorstellen, dass die eben vorgelesenen Kernzeitschriften relevantere Arbeiten zum Thema publizieren als die übrigen in SOLIS ausgewerteten Zeitschriften?

Die Frage 3 „Relevanz der Kernzeitschriften“ fokussiert auf die Einschätzung der Relevanz der Kernzeitschriften gegenüber den übrigen Zeitschriften. Auf diese Frage hat die Mehrheit (15 Topics) mit „ja“ geantwortet. Die Mehrheit der Befragten (N=9) glaubt, dass die Kernzeitschriften relevantere Arbeiten publizieren. Lediglich bei 5 Topics wird mit „nein“ geantwortet; hier glauben die Bearbeiter nicht an einen Relevanzvorteil dieser Zeitschriften. Bei 7 Topics wurde keine Meinung („weiß nicht“) zu dieser Frage geäußert. Folgerichtig ist, dass Personen, die spontan keine Kernzeitschriften nennen konnten (Frage 1), bei Frage 3 auch nicht von einem Relevanzvorteil dieser Zeitschriften ausgehen. Immerhin 6 Befragte (7 Topics) glauben, ohne spontan Kernzeitschriften benennen zu können (Frage 1), trotzdem an den Relevanzvorteil dieser Zeitschriften (siehe Tabelle 47).

Tabelle 47: Ergebnisse Frage 3. N=24 Topics

| Antworten zu Frage 3 | |
|----------------------------------|--------|
| „Relevanz der Kernzeitschriften“ | Topics |
| ja | 15 |
| nein | 5 |
| weiß nicht | 7 |

Die letzte Frage zielte auf weitere Zeitschriften, die nicht explizit in der Liste der Kernzeitschriften genannt wurden, aber dennoch regelmäßig Literatur zum behandelten Thema publizieren.

Frage 4: Welche Zeitschriften, die regelmäßig Literatur zu den obigen Topics beinhalten, vermissen Sie in dieser Liste? Zählen Sie bitte mehrere Zeitschriftentitel auf!

Tabelle 48: Ergebnisse Frage 4. N=24 Topics

| Antworten zu Frage 4 | |
|-------------------------|--------|
| „Weitere Zeitschriften“ | Topics |
| keine | 14 |
| 1 Zeitschrift | 6 |
| mehrere Zeitschriften | 4 |

Insgesamt konnten zu 14 Topics keine weiteren Zeitschriften genannt werden. Zu 10 Topics konnten eine (N=6) oder mehrere weitere Zeitschriften (N=4) genannt werden (siehe Tabelle 48).

Zum Schluss des Interviews wurden die Befragten ermuntert, Kommentare und Hinweise zum Interview abzugeben.

Die Ergebnisse der Interviews und die Kommentare werden im folgenden zusammenfassenden Abschnitt aufgegriffen.

7.3.2 Zusammenfassung

Folgende Diskussionsstränge ergeben die Resultate der Interviews:

Ein Erkenntnisgewinn der befragten Personen durch das Verlesen der Kernzeitschriften kann verdeutlicht werden. Für mehr als 50 % der Topics können die „Informationsexperten“ „ihre“ Kernzeitschriften nicht spontan benennen. Dieses Ergebnis überrascht; vermutlich sind aber die Zuordnungsprobleme in der Spezifität oder Interdisziplinarität der einzelnen Topics, der Erfahrungen auf dem entsprechenden Gebiet sowie der „künstlichen“ Fragesituation begründet. Es ist scheinbar schwierig, als nicht aktiv Forschende(r) die Kernzeitschriften eines engen bzw. übergreifenden Gebiets spontan zu benennen. Viele Interviewpartner erkennen „ihre“ Kernzeitschriften nach dem Vorlesen wieder und sind positiv überrascht von der Analyse (vgl. These 3 in Kapitel 10). Die Überraschung der Befragten beim Vorlesen der Zeitschriften und deren Zeitschriftenrängen war recht auffällig und demonstriert einen positiven Effekt der Kernzeitschriften-Analyse (Bradfordizing).

Der folgende Abschnitt fasst die Ergebnisse der Relevanzauswertungen zusammen.

7.4 Effekte von Bradfordizing beim IR

Verbessert das Bradfordizing den Anteil der relevanten Dokumente im ersten Drittel (Core) einer Dokumentmenge gegenüber dem zweiten und dritten Drittel (Zone 2 bzw. Zone 3)?

Lässt sich das Verfahren Bradfordizing als Re-Ranking-Ansatz auch erfolgreich auf andere Dokumenttypen übertragen?

Im folgenden Abschnitt wird untersucht, welche Auswirkungen sich beim Re-Ranking nach Bradfordizing auf die Relevanzverteilung im neuorganisierten Ergebnis beobachten lassen. Die Studie untersucht dazu standardisierte und nach Relevanz bewertete Fragestellungen nach dem Bradfordizing (Re-Ranking). Der Einsatz von Bradfordizing soll das Suchergebnis komplett neu zusammensetzen und es möglichst signifikant verbessern. Ziel ist es, mehr relevante Dokumente für das Topic im ersten Drittel der Ergebnismenge zu listen. Die Anwendung von Bradfordizing bei der Suche stellt somit einen Mehrwert dar: Bradfordizing soll den Anteil der relevanten Dokumente im vorderen Trefferlistenbereich erhöhen und insgesamt den Suchprozess durch die neue Struktur des Suchergebnisses positiv beeinflussen. Im Mittelpunkt der Analyse steht der Vergleich der durchschnittlichen Precision-Werte der drei typischen Dokumentzonen (Bradford-Zonen) für mehrere Dokumentenpools. Im Experiment wurden die beiden Dokumenttypen Zeitschriftenartikel und Monographien pro Topic untersucht.

Nachfolgend werden die Auswertungsergebnisse der beiden relevanzbewerteten Datensätze (CLEF und KoMoHe) präsentiert. Der Datensatz CLEF umfasst insgesamt 125 Topics der Jahrgänge 2003 bis 2007 und basiert auf sozialwissenschaftlichen Dokumenten des GIRT-Korpus und deren Relevanzbewertungen. Der Datensatz KoMoHe umfasst insgesamt 39 Topics der Information Retrieval-Evaluation im Projekt KoMoHe und basiert auf Dokumenten der Datenbanken SOLIS, SoLit, USB Köln OPAC, Psynindex, World Affairs Online, Medline. Alle Relevanzbewertungen der KoMoHe-Topics sind im Jahr 2007 erfolgt und lassen sich drei Testreihen (vgl. Abschnitt 5.4.2) zuordnen.

7.4.1 Precision innerhalb der Bradford-Zonen

Die folgenden Tabellen und Abbildungen zeigen die durchschnittlichen Precision-Werte der Bradford-Zonen (Core, Zone 2 bzw. Z2 und Zone 3 bzw. Z3) für die Topics der beiden Datensätze CLEF und KoMoHe.

Die beiden folgenden Abschnitte präsentieren die Precision-Werte und die Precision-Unterschiede zwischen den Zonen für Zeitschriftenartikel (article) und Monographien (monographs). Die Auswertungen werden jeweils getrennt voneinander dargestellt. Die Tabellen enthalten jeweils die Mittelwerte der Auszählungen für eine Topic-Reihe.

7.4.2 Zeitschriftenartikel

Auswertung der CLEF-Topics

Tabelle 49 zeigt die gemittelten Precision-Werte für 125 CLEF-Topics. Im Durchschnitt beträgt die Precision in der Kernzone 0,28. Die Precision sinkt in den folgenden Zonen

deutlich ab (Zone 2 = 0,23 und Zone 3 = 0,19). Die baseline-Precision fasst die durchschnittliche Precision über alle Zeitschriftenartikel zusammen.

Tabelle 49: Durchschnittliche Precision für Zeitschriftenartikel in den Bradford-Zonen (Core, Z2 und Z3) und baseline-Precision im CLEF-Datensatz (N=125 Topics)

| CLEF Article | Topics | P core | P Z2 | P Z3 | P baseline |
|--------------|--------|--------|-------|-------|------------|
| 2003 | 25 | 0,294 | 0,218 | 0,157 | 0,221 |
| 2004 | 25 | 0,226 | 0,185 | 0,134 | 0,179 |
| 2005 | 25 | 0,310 | 0,240 | 0,174 | 0,239 |
| 2006 | 25 | 0,288 | 0,267 | 0,244 | 0,265 |
| 2007 | 25 | 0,278 | 0,256 | 0,217 | 0,248 |

Die folgenden Abbildungen (Abbildung 7-16 bis Abbildung 7-20) visualisieren die Precision-Unterschiede der beiden Dokumentzonen Core und Zone 3 sowie die durchschnittliche Precision über alle Dokumente/Zeitschriftenartikel (baseline) des jeweiligen Topics. Die Precision-Werte sind jeweils nach der Precision im Core absteigend sortiert.

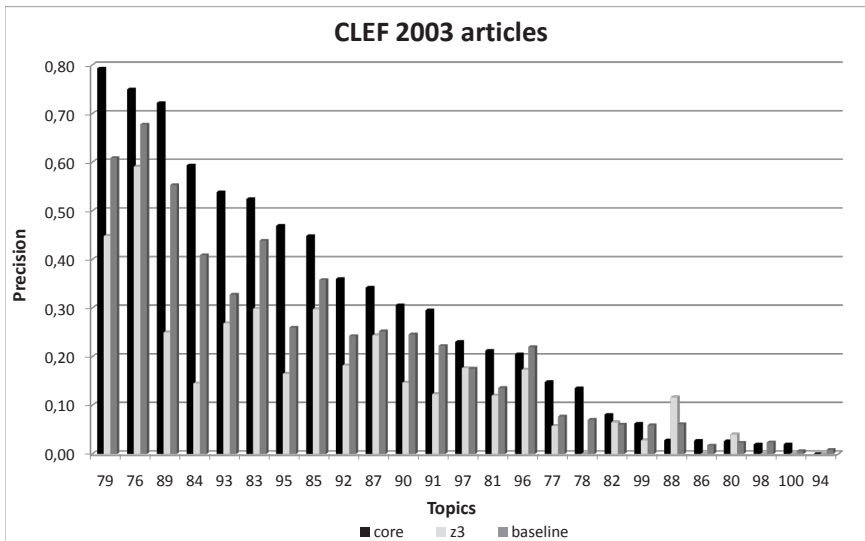


Abbildung 7-16: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2003.

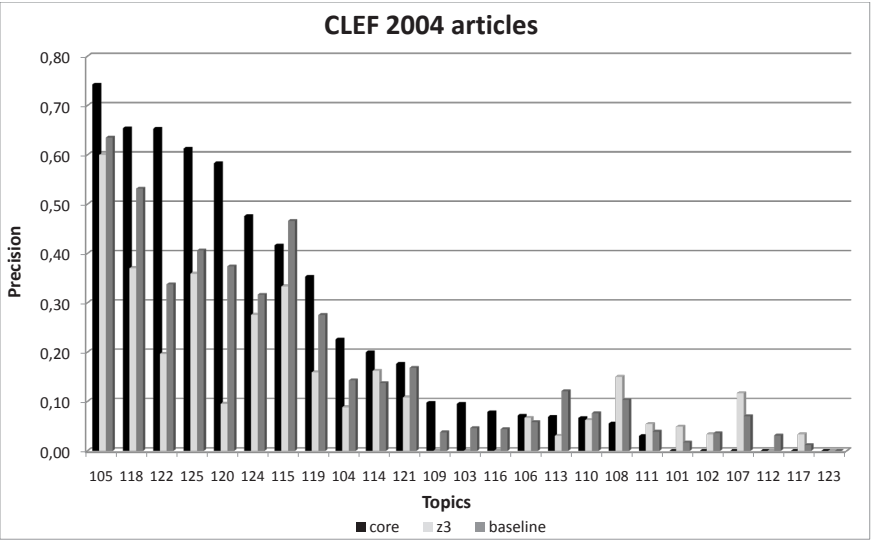


Abbildung 7-17: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2004.

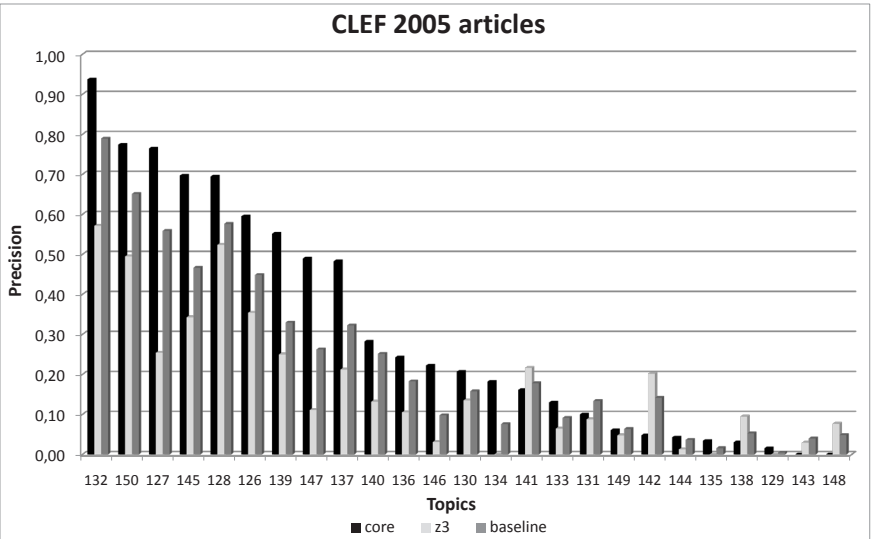


Abbildung 7-18: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2005.

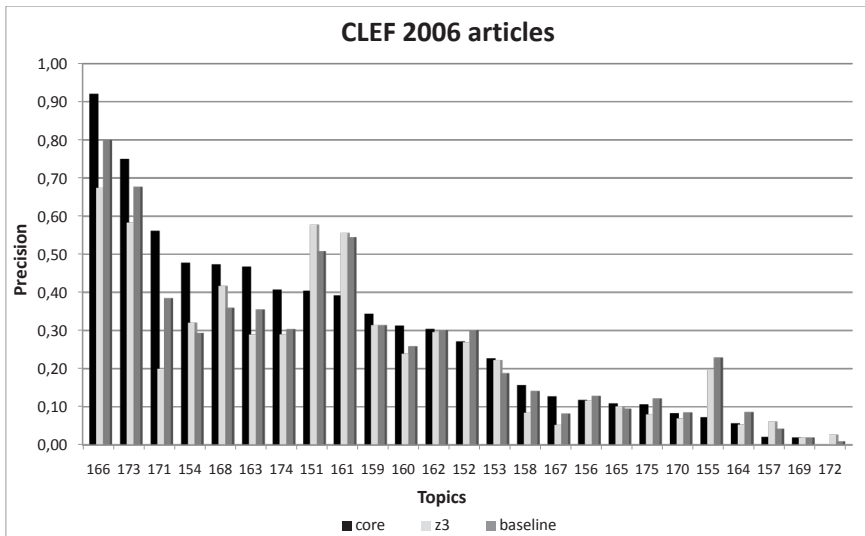


Abbildung 7-19: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2006.

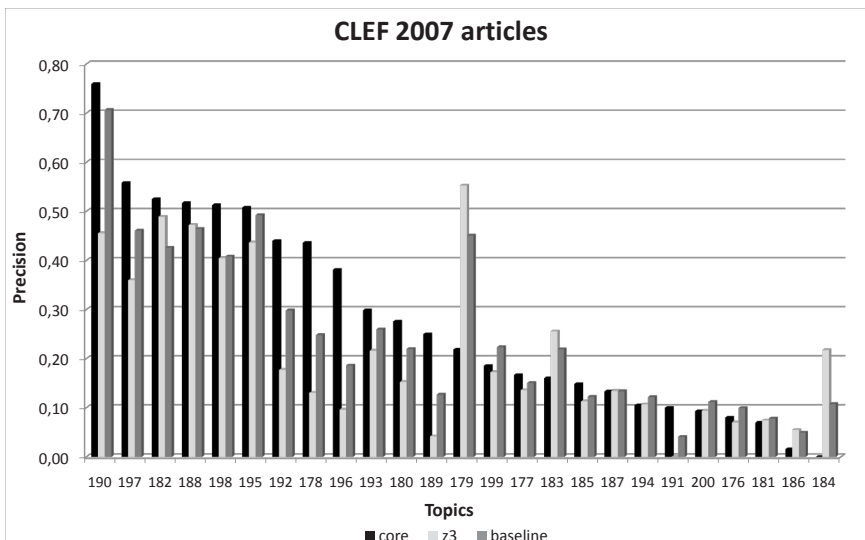


Abbildung 7-20: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel der 25 Topics aus dem Jahre 2007.

Tabelle 50 zeigt die prozentuale Verbesserung der Precision der drei Zonen untereinander. Als Vergleichswert ist zusätzlich die baseline-Precision mitgeführt. Signifikante Verbesserungen sind jeweils durch das Sternsymbol gekennzeichnet (*). Es wird nicht zwischen unterschiedlichen Signifikanzniveaus unterschieden. Die Signifikanztests (T-Test und Wilcoxon-Test) der jeweiligen Datensätze finden sich im Anhang „Signifikanztests“. Die Kernzone (core) enthält demnach über alle CLEF-Jahrgänge ca. 56 % mehr relevante Dokumente als Zone 3. Die Kernzone beinhaltet ebenfalls ca. 20 % mehr relevante Dokumente als die nachfolgende Zone 2. Zone 2 beinhaltet wiederum ca. 28 % mehr relevante Dokumente als Zone 3. Der Vergleich von Core und baseline-Precision fällt ebenfalls positiv für die Dokumente im Core aus. Die Zeitschriftenartikel im Core erzielen eine um knapp 22 % bessere Precision als der Durchschnitt über alle Zeitschriftenartikel.

Tabelle 50: Durchschnittliche Verbesserung der Precision für Zeitschriftenartikel zwischen den Bradford-Zonen (Core, Z2 und Z3) im CLEF-Datensatz

| CLEF article | Verbesserung core zu Z3 in % | Verbesserung core zu Z2 in % | Verbesserung Z2 zu Z3 in % | Verbesserung core zu baseline in % |
|--------------|---------------------------------|---------------------------------|-------------------------------|---------------------------------------|
| 2003 | 86,56 (*) | 34,57 (*) | 38,63 (*) | 32,65 (*) |
| 2004 | 69,23 (*) | 22,45 | 38,20 | 26,25 (*) |
| 2005 | 78,03 (*) | 29,05 (*) | 37,95 (*) | 29,52 (*) |
| 2006 | 17,63 | 7,66 | 9,27 | 8,46 |
| 2007 | 28,18 (*) | 8,31 | 18,35 | 11,77 |
| | 55,93 (*) | 20,41 (*) | 28,48 (*) | 21,73 (*) |

Auswertung der KoMoHe-Topics

Ein vergleichbares Ergebnis liefert der KoMoHe-Datensatz (siehe Tabelle 51). Die 39 Topics zeigen die gleiche Tendenz: Die Dokumente der Kernzone sind signifikant häufiger relevant als Dokumente der Zone 2 und Zone 3. Im Durchschnitt beträgt die Precision in der Kernzone 0,40. Die Precision sinkt in den folgenden Zonen ab (Zone 2 = 0,37 und Zone 3 = 0,34).

Tabelle 51: Durchschnittliche Precision für Zeitschriftenartikel in den Bradford-Zonen (Core, Z2 und Z3) im KoMoHe-Datensatz (N=39 Topics)

| KoMoHe article | Topics | P core | P Z2 | P Z3 | P baseline |
|----------------|--------|-----------|---------|---------|---------------|
| Test1 | 15 | 0,292 | 0,261 | 0,245 | 0,265 |
| Test2 | 12 | 0,215 | 0,202 | 0,192 | 0,202 |
| Test3 | 12 | 0,700 | 0,644 | 0,587 | 0,642 |

Die folgenden Abbildungen (Abbildung 7-21 bis Abbildung 7-23) visualisieren die Precision-Unterschiede der beiden Dokumentzonen Core und Zone 3 (z3) sowie die durchschnittliche Precision über alle Dokumente/Zeitschriftenartikel (baseline) des jeweiligen Topics. Die Precision-Werte sind jeweils nach der Precision im Core absteigend sortiert.

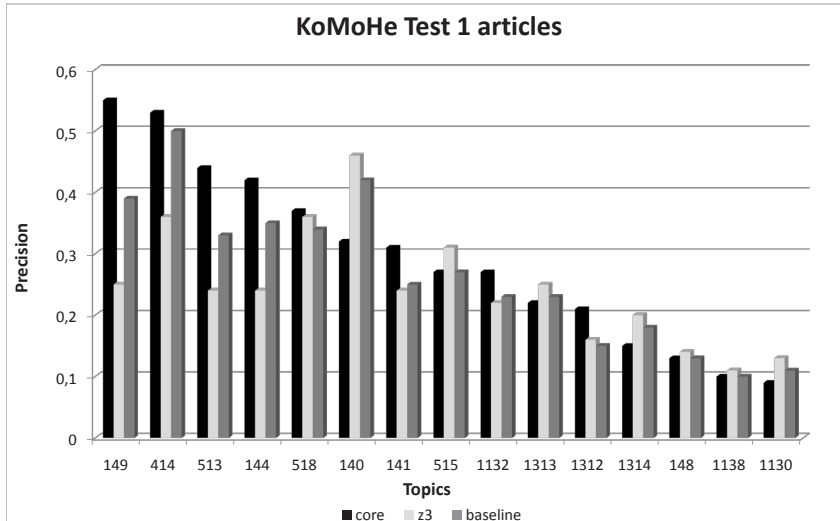


Abbildung 7-21: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel von 15 Topics des KoMoHe-Tests 1 (innerdisziplinärer Test).

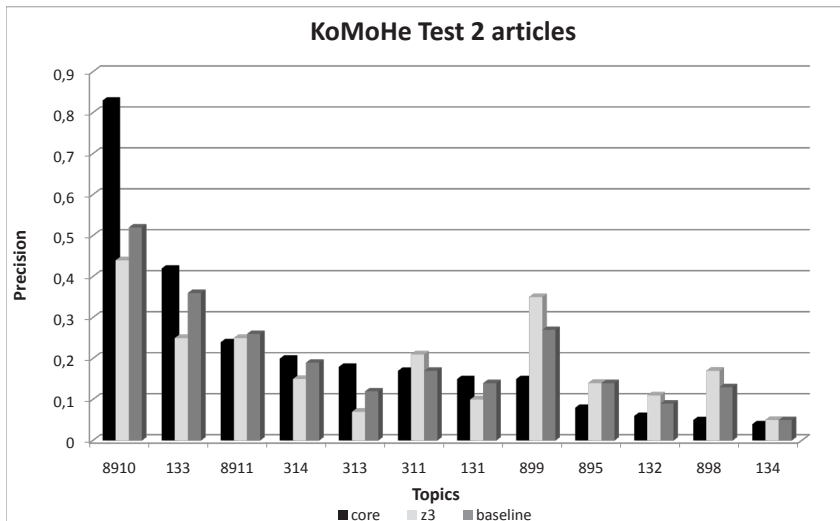


Abbildung 7-22: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel von 12 Topics des KoMoHe-Tests 2 (interdisziplinärer Test).

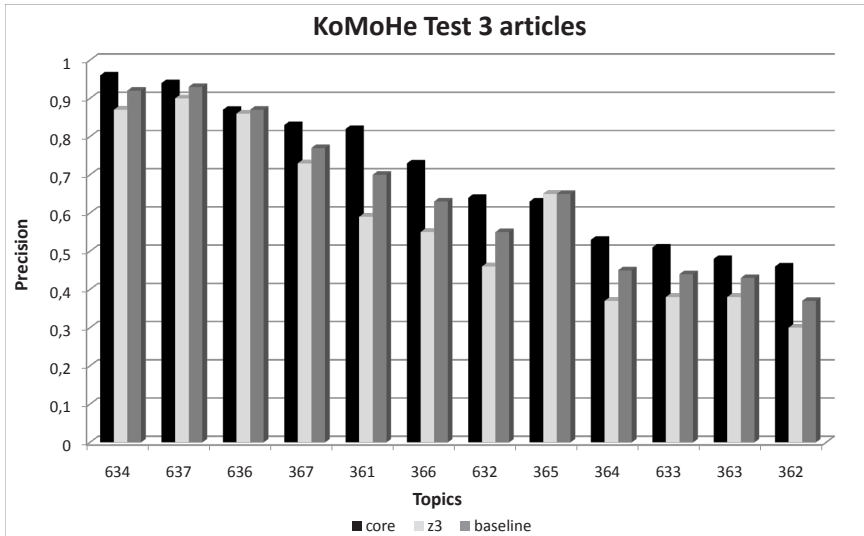


Abbildung 7-23: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Zeitschriftenartikel von 12 Topics des KoMoHe-Tests 3 (interdisziplinärer Test).

Tabelle 52 zeigt die prozentuale Verbesserung der Precision der drei Zonen untereinander. Als Vergleichswert ist zusätzlich die baseline-Precision mitgeführt. Signifikante Verbesserungen sind jeweils durch das Sternsymbol gekennzeichnet (*). Es wird nicht zwischen unterschiedlichen Signifikanzniveaus unterschieden. Die Signifikanztests (T-Test und Wilcoxon-Test) der jeweiligen Datensätze finden sich im Anhang „Signifikanztest KoMoHe (Zeitschriftenartikel)“. Die Kernzone (core) enthält demnach über alle KoMoHe-Tests ca. 17 % mehr relevante Dokumente als Zone 3. Die Kernzone beinhaltet ebenfalls ca. 9 % mehr relevante Dokumente als die nachfolgende Zone 2. Zone 2 beinhaltet wiederum ca. 7 % mehr relevante Dokumente als Zone 3. Der Vergleich von Core und baseline-Precision fällt ebenfalls positiv für die Dokumente im Core aus. Die Zeitschriftenartikel im Core erzielen eine um knapp 8 % bessere Precision als der Durchschnitt über alle Zeitschriftenartikel.

Beim Vergleich dieser Ergebnisse zeigt sich, dass die Verbesserung der Precision zwischen den Zonen bei den KoMoHe-Tests nicht so deutlich ausfällt wie bei den CLEF-Topics.

Tabelle 52: Durchschnittliche Verbesserung der Precision für Zeitschriftenartikel zwischen den Bradford-Zonen (Core, Z2 und Z3) im KoMoHe-Datensatz

| KoMoHe article | Verbesserung core zu Z3 in % | Verbesserung core zu Z2 in % | Verbesserung Z2 zu Z3 in % | Verbesserung core zu baseline in % |
|----------------|---------------------------------|---------------------------------|-------------------------------|---------------------------------------|
| Test1 | 18,82 | 11,75 | 6,32 | 9,84 |
| Test2 | 11,58 | 6,16 | 5,11 | 6,12 |
| Test3 | 19,32 (*) | 8,67 (*) | 9,80 (*) | 9,00 (*) |
| | 16,57 (*) | 8,86 | 7,08 (*) | 8,32 (*) |

7.4.3 Monographien

Auswertung der CLEF-Topics

Zum Vergleich findet sich die Auswertung der Precision auch für den Dokumenttyp Monographie (siehe Tabelle 53). Aus dem CLEF-Datensatz wurden drei Jahrgänge analysiert. Der Precision-Vorteil der Kernzone gegenüber den nachfolgenden Zonen fällt bei den Monographien bei den CLEF-Topics geringer aus.

Tabelle 53: Durchschnittliche Precision für Monographien in den Bradford-Zonen (Core, Z2 und Z3) im CLEF-Datensatz (N=75 Topics)

| | N | P core | P Z2 | P Z3 | P baseline |
|------------|----|-----------|---------|---------|---------------|
| CLEF books | | | | | |
| 2005 | 25 | 0,205 | 0,162 | 0,188 | 0,188 |
| 2006 | 25 | 0,284 | 0,283 | 0,236 | 0,251 |
| 2007 | 25 | 0,238 | 0,205 | 0,234 | 0,227 |

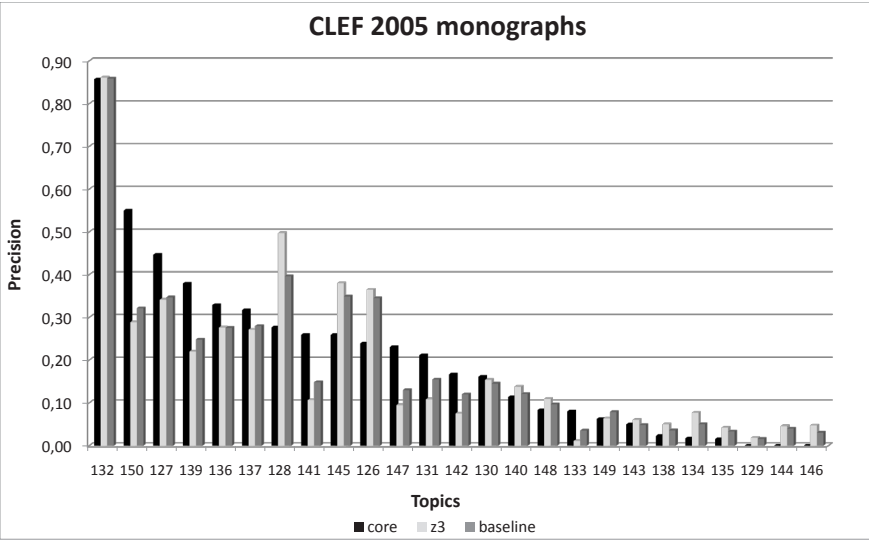


Abbildung 7-24: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Monographien der 25 Topics aus dem Jahre 2005.

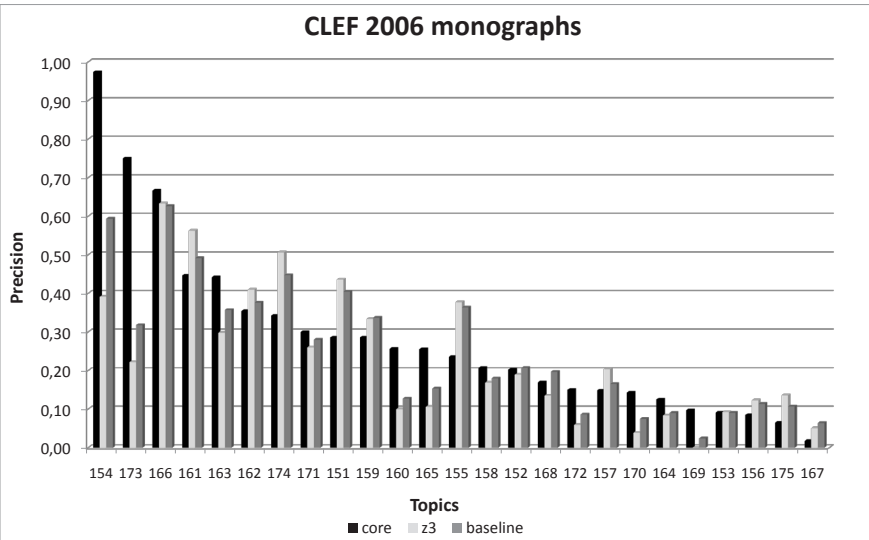


Abbildung 7-25: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Monographien der 25 Topics aus dem Jahre 2006.

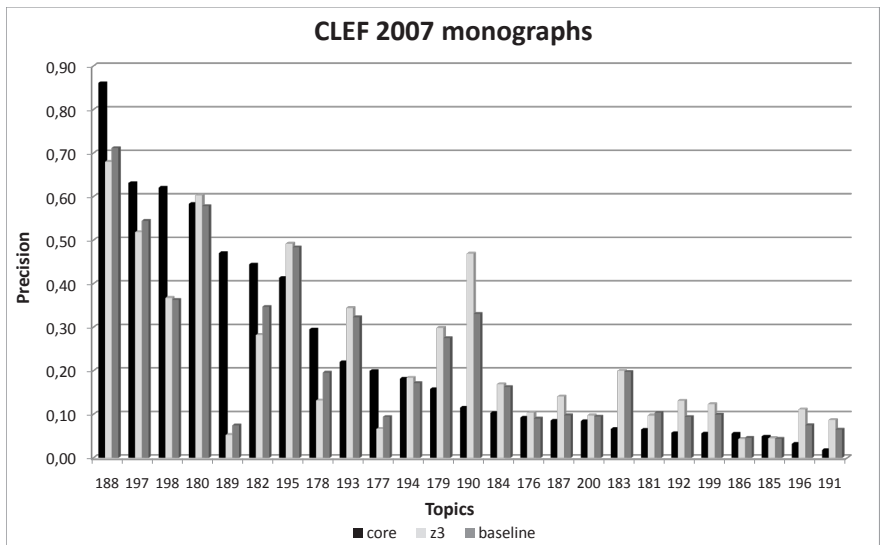


Abbildung 7-26: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Monographien der 25 Topics aus dem Jahre 2007.

Tabelle 54 zeigt die prozentuale Verbesserung der Precision der drei Zonen untereinander. Als Vergleichswert ist zusätzlich die baseline-Precision mitgeführt. Signifikante Verbesserungen finden sich nicht bei den Ergebnissen. Die Kernzone (core) enthält demnach über alle drei CLEF-Jahrgänge ca. 10 % mehr relevante Dokumente als Zone 3. Die Kernzone beinhaltet ebenfalls ca. 14 % mehr relevante Dokumente als die nachfolgende Zone 2. Zone 2 beinhaltet um ca. 2 % weniger relevante Dokumente als Zone 3. Der Vergleich von Core und baseline-Precision fällt ebenfalls positiv für die Dokumente im Core aus. Die Zeitschriftenartikel im Core erzielen eine um knapp 9 % bessere Precision als der Durchschnitt über alle Zeitschriftenartikel.

Tabelle 54: Durchschnittliche Verbesserung der Precision für Monographien zwischen den Bradford-Zonen (Core, Z2 und Z3) im CLEF-Datensatz

| CLEF books | Verbesserung core zu Z3 in % | Verbesserung core zu Z2 in % | Verbesserung Z2 zu Z3 in % | Verbesserung core zu baseline in % |
|------------|------------------------------|------------------------------|----------------------------|------------------------------------|
| 2005 | 9,09 | 26,69 | -13,89 | 8,98 |
| 2006 | 19,98 | 0,33 | 19,58 | 13,07 |
| 2007 | 2,09 | 16,13 | -12,09 | 5,21 |
| | 10,39 | 14,38 | -2,13 | 9,09 |

Die Precision für 75 getestete Topics in der Datenbank SOLIS erhöht sich nach Bradfordizing der Monographien um durchschnittlich 10 % (Vergleich der Precision der Kernverlage und der Verlage in Zone 3) (nicht signifikant).

Auswertung der KoMoHe-Topics

Ein ähnliches Bild zeigt sich auch bei den KoMoHe-Topics für den Dokumenttyp Monographie (siehe Tabelle 55). **Der Precision-Vorteil der Kernzone gegenüber den nachfolgenden Zonen ist ebenfalls sichtbar und mit den Precision-Werten der Zeitschriftenartikel vergleichbar.**

Tabelle 55: Durchschnittliche Precision für Monographien in den Bradford-Zonen (Core, Z2 und Z3) im KoMoHe-Datensatz (N=30 Topics)

| KoMoHe books | N | P core | P Z2 | P Z3 | P baseline |
|--------------|----|-----------|---------|---------|---------------|
| Test1 | 10 | 0,247 | 0,181 | 0,188 | 0,195 |
| Test2 | 15 | 0,255 | 0,251 | 0,243 | 0,244 |
| Test3 | 5 | 0,684 | 0,612 | 0,617 | 0,637 |

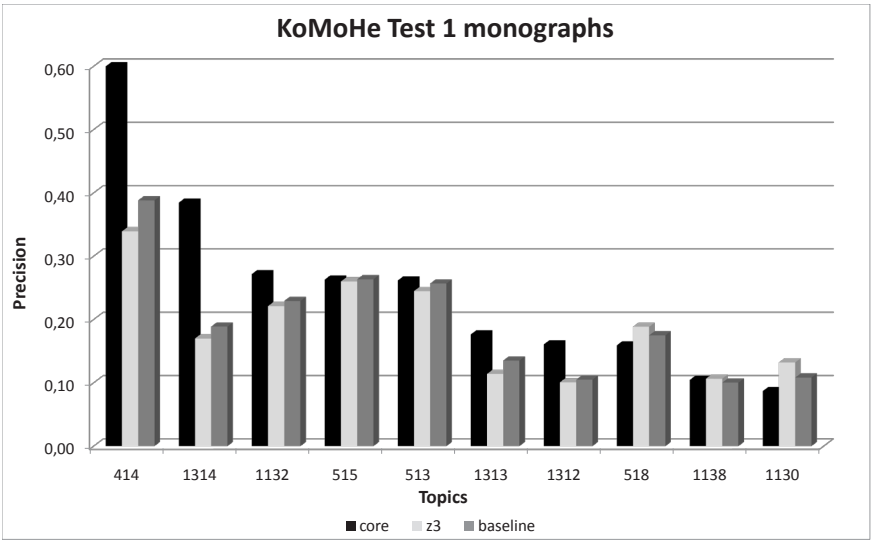


Abbildung 7-27: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Monographien von 10 Topics des KoMoHe-Tests 1 (innerdisziplinärer Test).

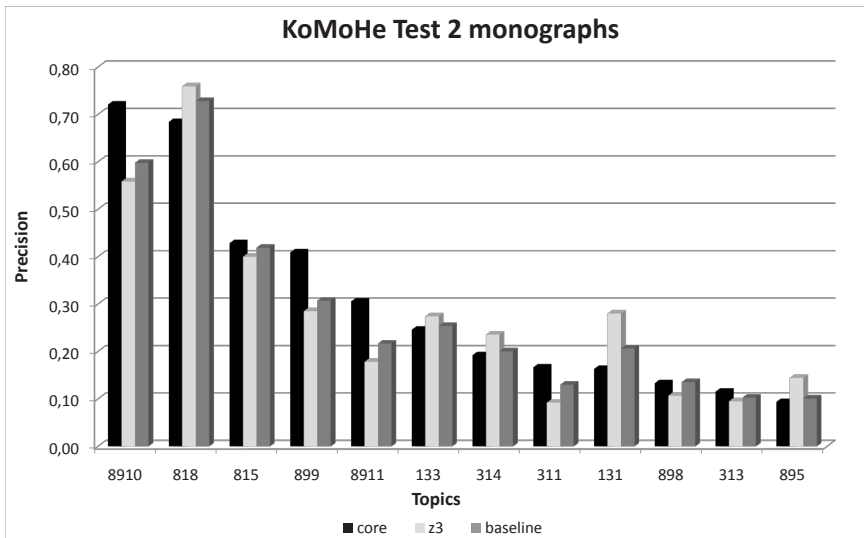


Abbildung 7-28: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Monographien von 15 Topics des KoMoHe-Tests 2 (interdisziplinärer Test).

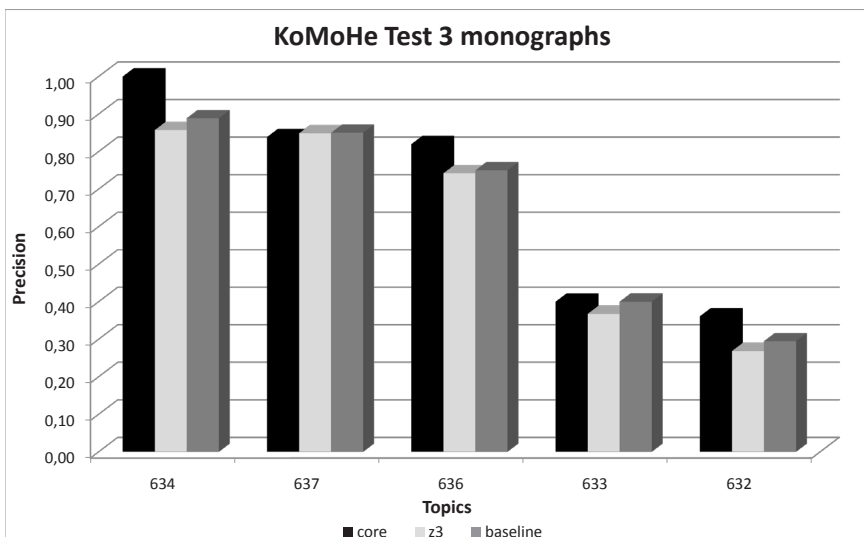


Abbildung 7-29: Verteilung der Precision-Werte je Topic für die Dokumentzonen (Core, Zone 3) sowie den Durchschnitt über alle Dokumente (baseline). Die Precision-Werte sind nach der Precision im Core absteigend sortiert. Analysiert wurden die Monographien von 5 Topics des KoMoHe-Tests 3 (interdisziplinärer Test).

Tabelle 56 zeigt die prozentuale Verbesserung der Precision der drei Zonen untereinander. Als Vergleichswert ist zusätzlich die baseline-Precision mit aufgeführt. Die Kernzone (core) enthält demnach über alle drei KoMoHe-Tests ca. 16 % mehr relevante Dokumente als Zone 3. Diese Verbesserung ist signifikant. Die Kernzone beinhaltet ebenfalls ca. 17 % mehr relevante Dokumente als die nachfolgende Zone 2. Zone 2 beinhaltet in etwa gleich viele relevante Dokumente wie Zone 3. Der Vergleich von Core und baseline-Precision fällt ebenfalls positiv für die Dokumente im Core aus. Die Zeitschriftenartikel im Core erzielen eine um knapp 13 % bessere Precision als der Durchschnitt über alle Zeitschriftenartikel. Diese Verbesserung ist ebenfalls signifikant.

Tabelle 56: Durchschnittliche Verbesserung der Precision für Monographien zwischen den Bradford-Zonen (Core, Z2 und Z3) im KoMoHe-Datensatz

| KoMoHe books | Verbesserung core zu Z3 in % | Verbesserung core zu Z2 in % | Verbesserung Z2 zu Z3 in % | Verbesserung core zu baseline in % |
|--------------|---------------------------------|---------------------------------|-------------------------------|---------------------------------------|
| Test1 | 31,51 | 36,27 | -3,50 | 26,58 |
| Test2 | 4,80 | 1,57 | 3,17 | 4,56 |
| Test3 | 10,91 | 11,83 | -0,82 | 7,41 |
| | 15,74 (*) | 16,56 | -0,38 | 12,85 (*) |

7.4.4 Zusammenfassung

Die Precision für 125 getestete Topics in der Datenbank SOLIS erhöht sich nach Bradfordizing um durchschnittlich 56 % (Vergleich der Precision der Kernzeitschriften und der Zone 3-Zeitschriften). Diese signifikante Verbesserung durch Bradfordizing tritt auch für andere Datenbanken und Disziplinen ein. Die Precision für 39 getestete Topics in den Datenbanken SOLIS, SoLit, USB Köln OPAC, Psyndex, Econis, World Affairs Online, Medline erhöht sich nach Bradfordizing um durchschnittlich 17 % (Vergleich der Precision der Kernzeitschriften und der Zone 3-Zeitschriften).

Die Dokumente der Kernzone zeigen für die meisten Testreihen signifikant höhere Precision als Dokumente der Zone 2 und Zone 3 (siehe dazu auch T-Tests und Wilcoxon-Test im Anhang „Signifikanztests“). Sowohl für die Zeitschriften als auch für die Monographien kann diese Aussage auf einer sehr breiten Basis von Themen und Fragestellungen an zwei unabhängigen Dokumentkorpora empirisch nachgewiesen werden. Das Verfahren Bradfordizing bzw. das Re-Ranking nach der Zeitschriftenhäufigkeit ist demnach für die erfasste Literatur in Fachdatenbanken (insbesondere Zeitschriftenliteratur) zu einem Thema aufgrund der signifikanten Precision-Verbesserung sehr vielversprechend.

Nachdem in diesem Kapitel die Ergebnisse der Untersuchungen (aus Abschnitt 5) beschrieben wurden, sollen im folgenden Kapitel die Hauptergebnisse diskutiert werden.

8 Diskussion

Das folgende Kapitel greift die Ergebnisse des vorherigen Kapitels auf und diskutiert anhand der Fachliteratur die Hauptergebnisse der Untersuchung. Die Forschungsfragen werden der Diskussion jeweils vorangestellt. Die Diskussion der einzelnen Abschnitte beginnt jeweils mit der Auseinandersetzung der konkreten Untersuchungsergebnisse (siehe Kapitel 7) dieser Studie. Angrenzende Fragestellungen und Fachliteratur werden im Anschluss diskutiert.

8.1 Effekte von Crosskonkordanzen für das IR

Welche Auswirkungen und Effekte hat das Konzept semantischer Heterogenitätsbehandlung am Beispiel Crosskonkordanzen für die Suche nach Literatur?

Der folgende Abschnitt bezieht sich auf die Ergebnisse in Abschnitt 7.1.

Die Evaluation der Retrieval-Effektivität von Crosskonkordanzen stellt bis dato ein Forschungsdesiderat dar. In den aktuellen Reviews zum Stand der Forschung bei Terminologie-Diensten ((Patel et al., 2005, Tudhope et al., 2006, Zeng und Chan, 2004)) wird die Bedeutung dieses Themas zwar erwähnt, es finden sich aber nach Kenntnisstand des Autors keine standardisierten Retrieval-Untersuchungen, die empirisch die Effekte von Terminologie-Mappings beim IR untersuchen und eventuelle IR-Mehrwerte evaluieren.

Die Evaluation der Effektivität der Crosskonkordanzen für die Suche in heterogen erschlossenen Datenbanken über ein klassisches IR-Testverfahren (siehe Abschnitt 5.2.1) verdeutlicht, dass der Einsatz von Crosskonkordanzen die IR-Werte Recall und Precision signifikant verbessern kann. Insbesondere für die Suche mit kontrollierten Termen (gültigen Deskriptoren) kann der Einsatz von Crosskonkordanzen die beiden Standardmaße im IR positiv beeinflussen und insgesamt zu einer Erhöhung von Recall und Precision führen (siehe Tabelle 34 bis Tabelle 37). Für natürlichsprachige Terme (Test B in Abschnitt 7.1.2) kommt es lediglich zu einer Erhöhung des Recall (siehe Tabelle 38 bis Tabelle 40). Die Ergebnisse der Experimente mit der Freitext-Suche könnten wahrscheinlich sehr verbessert werden, wenn die kontrollierten Terme und die natürlichsprachigen Terme besser bei der Anfrageformulierung integriert werden würden, anstatt sie nur aneinanderzufügen. Eine Möglichkeit wäre, die natürlichsprachigen Terme (automatisch) a priori in kontrollierte Terme zu übersetzen, damit eine größere Chance besteht, unterschiedliche kontrollierte Vokabulare und Datenbanken aufeinander abzubilden (siehe Ansatz bei (Petras, 2006)). Das Retrieval mit Benutzertermen steht nicht im Fokus der Untersuchung und soll in künftigen IR-Untersuchungen (z. B. im GESIS-Projekt IR-Mehrwertdienste ab Januar 2009) evaluiert werden (siehe dazu Kapitel 9).

Der stärkste Effekt tritt bei Crosskonkordanzen auf, die Erschließungssprachen involvieren, die in unterschiedlichen Disziplinen verortet sind (siehe Tabelle 37). **Diese interdisziplinären Crosskonkordanzen** (siehe Term-Beispiele im Anhang „Ausschnitte aus der Crosskonkordanz Thesaurus Sozialwissenschaften (TheSoz) zu Standard Thesaurus Wirtschaft (STW)“ und „Ausschnitte aus der Crosskonkordanz Psynindex Terms zu Medical Subject Headings (MeSH)“) verbinden im IR-Test beispielsweise Datenbanken der Medi-

zin mit Psychologie (und umgekehrt Psychologie mit Medizin) oder Sozialwissenschaften mit Wirtschaftswissenschaften (und umgekehrt Wirtschaftswissenschaften mit Sozialwissenschaften). Die disziplinübergreifende Suche mit kontrollierten Termen scheint damit für den Nutzer ein sehr geeignetes Einsatzszenario für Terminologie-Dienste zu sein (siehe Abschnitt 3.2.5).

Die positiven Effekte beim Retrieval in unserem Szenario werden zusätzlich durch die Übersetzungsleistung der Crosskonkordanzen verstärkt (siehe dazu Beispiel in Abschnitt 6.1). Das Entfernen einer bilingualen Crosskonkordanz (Verbindungen zwischen deutschen und englischen Termen) aus dem Test (siehe Tabelle 35) **hat in unserer Untersuchung** zur Folge, dass die Verbesserung von Recall und Precision jeweils um etwa 10 % absinkt. Dieses Ergebnis verdeutlicht, dass insbesondere bei der Zusammenführung von mehrsprachigen Datenbanken und Erschließungssprachen Crosskonkordanzen einen erheblichen Mehrwert generieren. Die intellektuelle Übersetzungsleistung der Crosskonkordanz kann dabei nicht mit der rein automatischen Übersetzung von Termen⁸⁷ z. B. bei Internet-Suchmaschinen bzw. einfachen Query-Expansion-Verfahren (vgl. (Efthimiadis, 1996, Tudhope et al., 2006)) verglichen werden. Der Test A in Abschnitt 7.1.1 zeigt, dass die Reformulierung einer Deskriptoren-Anfrage durch Crosskonkordanzen in die korrekten kontrollierten Terme über Disziplin- und Sprachgrenzen hinweg die besten Retrieval-Effekte erbringt. Für künftige Mapping-Initiativen (siehe Übersicht in (Patel et al., 2005)) kann dieses Ergebnis u. U. Hinweise für die Auswahl der zu verbindenden kontrollierten Vokabulare geben.

Die theoretische Betrachtungsweise bei M. Doerr (2001), die als ideale Umgebung für die Erstellung von Crosskonkordanzen gleichartige kontrollierte Vokabulare innerhalb einer Domäne annimmt, um die Vagheit zwischen den Begrifflichkeiten zu reduzieren, existiert in unserem pragmatischen Untersuchungsszenario nicht.

„We further assume that the set of objects in all collections is basically of the same nature and from one domain, to exclude another source of vagueness.” (Doerr, 2001)

Im Gegenteil stellen wir gerade für Szenarien mit großer Vagheit (Bilingualität der verbundenen Vokabulare und Disziplinüberstieg) die größten Retrieval-Verbesserungen fest (siehe Tabelle 37). Während zwei Vokabulare aus dem gleichen Sprachbereich (z. B. Deutsch), die zudem aus einer Domäne stammen (z. B. Sozialwissenschaften) und ggf. vom gleichen Vokabulartyp sind (z. B. Thesaurus), also nach Doerr im Hinblick auf Wortverwendung und Abdeckung/Scope der Terme relativ ideale Mapping-Bedingungen aufweisen, schneiden gerade diese Crosskonkordanzen im IR-Test weniger gut ab (siehe intradisziplinäre Crosskonkordanzen). Gründe dafür sind vermutlich der hohe Anteil an syntaktisch zeichengleichen Termen in den verbundenen Vokabularen der intradisziplinären Crosskonkordanzen (Nulltransformationen bei (Marx, 2005)). Gerade diese Nulltransformationen haben keinen Effekt auf die Suche. Syntaktische Abgleiche (wie z. B. bei (Nikolai, 2002) beschrieben) zwischen zwei Vokabularen einer Sprache und innerhalb einer Domäne im Projekt KoMoHe zeigen, dass der Anteil der zeichengleichen Transfor-

87 <http://translate.google.com>

mationen (Nulltransformationen) deutlich höher ist als bei Crosskonkordanzen zwischen unterschiedlichen Domänen und Sprachen (siehe dazu Abschnitt 6.1).

“Another major issue for a growing terminology network is the scale and overlap of cross-concordances. The more vocabularies are mapped to each other, the more terms occur multiple times in variant mappings⁸⁸, which makes automatic query expansion more imprecise. On the other hand, the more vocabularies are added in such a network, the more inferences can be drawn for additional mappings. Indirect mappings via a pivot vocabulary could help in connecting vocabularies that haven’t been mapped to each other. A sufficiently large network could assist in reducing the mapping errors introduced by statistical or indirect mappings.” (Mayr und Petras, 2008)

Viele weitere Fragestellungen ergeben sich, wenn man ein Terminologie-Netz aus Crosskonkordanzen und anderen Mapping-Arten erstellt, z.B.: Wie viele Ausdrücke finden sich für ein Konzept? Welche Konzepte können verbunden werden? Sind die Vokabulare im Scope weiter oder enger? Welche Terminologien sind einander sehr ähnlich? Für welche Vokabulare sind automatische Mappings besonders lohnend? Welche Disziplinen/Fächer grenzen aneinander oder sind zu weit voneinander entfernt? Wie sehr überlappen sich unterschiedliche Datenbanken oder kontrollierte Vokabulare in einem bestimmten Fach? Wie kann das Zusammenspiel von automatischen und intellektuellen Verfahren zur Erstellung von Mappings gestaltet werden?

Gänzlich ausschließen lassen sich die semantischen Unterschiede (Semantik der Konzepte und Semantik der Beziehungen zwischen Konzepten) zwischen den kontrollierten Vokabularen durch die intellektuellen Verfahren der Heterogenitätsbehandlung auch bei idealen Bedingungen nicht (zur Problematik der semantischen Äquivalenz siehe die unterschiedlichen Äquivalenz-Beziehungen „partial equivalence“, „exact equivalence“, „inexact equivalence“, „single to multiple equivalence“ bei (Doerr, 2001)). Dies hat zur Folge, dass weitere Evaluationen die Qualität bzw. den Ballast (Recall/Precision) der zusätzlichen Dokumente durch die Termtransformationsverfahren auf empirischer Basis nachweisen müssen. Es gilt daher, zusätzliche Mechanismen zu konzipieren, die zu einem Ausgleich der durch die Heterogenitätsbehandlung entstandenen Unschärfe zwischen den Vokabularen (z.B. semantische Unterschiede der Deskriptoren und Indexierungspraxis) führen. Neben den Problemen bei der semantischen Überführung eines Vokabulars in ein anderes, die bei Doerr beschrieben werden (2001), ergeben sich insbesondere bei der praktischen Arbeit Probleme, die in den strukturellen Unterschieden der beteiligten Vokabulare begründet sind (vgl. (Hodge, 2000)). Beispielsweise ist es außerordentlich vage, einen Deskriptor aus einer unkommentierten Schlagwortliste, die keine Scope Notes und internen Verweisungen aufweist, mit einem „semantisch klaren“, also ausdefinierten Deskriptor eines Thesaurus zu verbinden (vgl. dazu (Schwarz und Umstätter, 1999) und deren Begriff des semiotischen Thesaurus). (Nikolai, 2002) schlägt in diesem

88 For example: term A from vocabulary 1 also occurs in vocabulary 2. A variant mapping exists when term A from vocabulary 1 is mapped to term B in vocabulary 3, but term A from vocabulary 2 is mapped to term C in vocabulary 3. This might be the correct mapping because the concepts in the different vocabularies are differently connotated but most of the time this will introduce noise to the network.

Zusammenhang vor, die Vokabulare vor der Integration (er spricht von Thesaurusföderationen) über unterschiedliche Metriken bzgl. ihrer Strukturen zu untersuchen (siehe dazu auch (Vizine-Goetz et al., 2004)).

Die Ergebnisse des IR-Tests zur Evaluation der Effektivität von Crosskonkordanzen sind sicher nur ein erster Schritt, um die Effekte von intra- und interdisziplinären Crosskonkordanzen aufzuzeigen. Nach dem heutigen Stand der Forschung liegen die Stärken der intellektuellen Crosskonkordanzen bei den sprachübergreifenden (bilingualen) und den interdisziplinären Mappings. Künftige Untersuchungen der Retrieval-Effektivität von Crosskonkordanzen sollen zudem den Mehrwert von mehreren integrierten Datenbanken (wie in Abbildung 3-8 dargestellt) evaluieren.

Wie erste Vergleiche zwischen intellektuellen und automatischen Verfahren zur Erstellung von Crosskonkordanzen im intradisziplinären Bereich zeigen, unterscheidet sich die Qualität der erstellten Term-Beziehungen nicht wesentlich voneinander (siehe z. B. Abbildung 8-1). **Die Untersuchung von Term-Recall und Term-Precision zwischen intellektuellen und automatischen Verfahren in einem monolingualen und intradisziplinären Szenario** (zwei deutschsprachige Vokabulare in den Sozialwissenschaften) bei (Marx, 2005) zeigt dies. Das Ergebnis erklärt sich z. T. durch den hohen Überlappungsgrad durch zeichengleiche Terme (Nulltransformationen) zwischen Thesaurus Sozialwissenschaften und Schlagwortnormdatei.

„Abschließend bleibt festzuhalten, dass intellektuelle und statistische Transformationen widererwartend eine sehr ähnliche Effektivität aufweisen.“ (Marx, 2005)

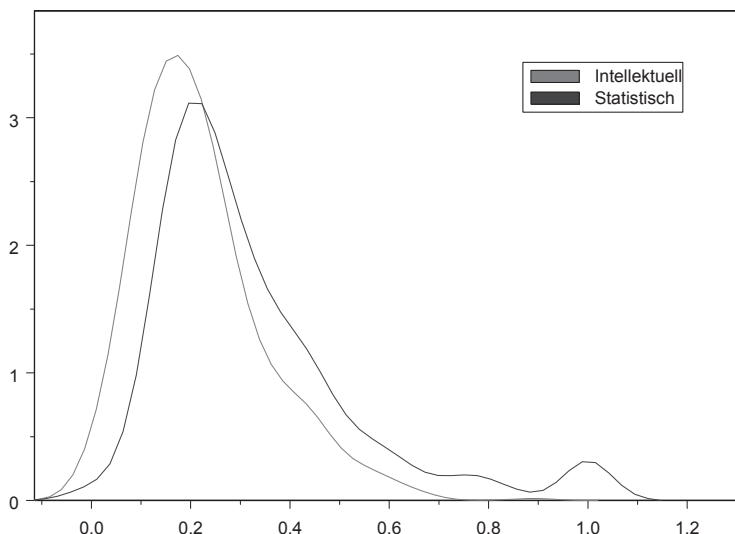


Abbildung 8-1: Vergleich Term-Precision intellektuell – statistisch, Korpus USB (aus Marx, 2005)

Die Ergebnisse bei Marx werden durch eine aktuelle Untersuchung im Bereich der Agrarwissenschaft ergänzt (siehe (Lauser et al., 2008)). Diese Untersuchung von 644 zufällig

ausgewählten, automatisch generierten Mappings zwischen AGROVOC und NALT ergibt eine Fehlerquote von ca. 14 % für semantisch nicht korrekte Verbindungen (siehe Abbildung 8-2). Die größten Schwierigkeiten zeigen die aktuellen automatischen Verfahren beim Generieren von hierarchischen Relationen sowie beim Ausschluss von Deskriptoren (vgl. Null-Relation in Abschnitt 3.2.2). Lauser et al. (2008) können exemplarisch zeigen, dass die automatischen Mapping-Verfahren insbesondere in allgemeineren Bereichen der Thesauri Schwächen aufweisen und häufiger zu falschen Mappings kommen.

“Things get potentially more difficult in the case of more diversified groups/categories (in our case just summarized as Miscellaneous). Here, often background knowledge is needed to infer the correct mapping, and automatic mapping tools are able to identify only very little of these correctly. Most of the automatic suggestions are simply wrong or should not be equivalence relationships but broader, narrower or related terms.” (Lauser et al., 2008)

Für den überwiegenden Teil der geographischen Terme, aber auch in Kategorien der biologischen und chemischen Fachterme erzielen die automatischen Verfahren gute Mapping-Ergebnisse.

| should be: | < | > | null (0) | ^ | total wrong |
|-----------------------|-----------|-----------|-----------|-----------|------------------|
| Taxonomic | 2.7% (7) | 0.38% (1) | 5.7% (15) | 0.38% (1) | 9.2% (24 of 262) |
| Biological / Chemical | 2.3% (2) | 1.2% (1) | 11% (9) | 0% (0) | 14% (12 of 84) |
| Miscellaneous | 1.4% (4) | 0.36% (1) | 14% (38) | 3.3% (9) | 19% (52 of 277) |
| all groups | 2.0% (13) | 0.0% (3) | 9.6% (62) | 1.5% (10) | 14% (88 of 643) |

Abbildung 8-2: Fehlerquote der automatisch erstellten Mappings zwischen AGROVOC und NALT (Daten aus (Lauser et al., 2008))

Die Autoren fassen zusammen: „The bottom line is that for the moment, mapping should not be seen as a monolithic exercise, but we can take the best of both approaches and use automatic mapping approaches to get to the simple and easy lexical mappings and then use human knowledge to control the ambiguous cases.” (Lauser et al., 2008)

Bezogen auf das koordinierte Zusammenspiel von automatischen und intellektuellen Verfahren zum Mapping von kontrollierten Vokabularen kommen (Hellweg et al., 2001) und Doerr zum gleichen Fazit.

„Statistical and neural network methods do not easily allow interpretation of the intellectual nature of a given link, if at all. They are far cheaper, however, and can detect relations of which humans are unaware. The ultimate precision is usually low. ... we are convinced that the future lies in the coordinated combination of intellectual and statistical methods ...“ (Doerr, 2001)

Es ist bekannt, dass bei Fachdatenbanken einzelne Zeitschriften durch die Indexierungs- bzw. die Auswertungspolitik des Datenbankproduzenten sehr unterschiedlich repräsentiert sind. Während bestimmte Zeitschriften in einer Datenbank zu 100 % ausgewertet werden (cover-to-cover), werden andere Zeitschriften deutlich selektiver indiziert und dafür u. U. in anderen Datenbanken vollständiger nachgewiesen. Diese Vorge-

hensweise ist keine Besonderheit, sondern gängige Praxis bei allen Datenbankanbietern (z. B. auch bei der Datenbank SOLIS, die in dieser Arbeit häufiger verwendet wurde). Wie bereits erwähnt, erfassen die großen angloamerikanischen Datenbanken (z. B. WoS) immer stärker nur den Kern von Bradfords bzw. Garfields Gesetz der Konzentration (Garfield, 1971, Garfield, 1990, Garfield, 1996). Eine aktuelle Untersuchung (siehe Abschnitt 6.4) zeigt diesen Trend exemplarisch an einer deutschsprachigen Zeitschrift (siehe (Mayr und Umstätter, 2008)). Vergleiche von Datenbanken haben bezüglich der dort auffindbaren bzw. sichtbaren Publikationen wiederholt Recall-Ratios von ca. 40 % erbracht. So kennen erfahrene Rechercheure das klassische Motto, dass eine Datenbank nicht ausreicht, um eine umfassende Recherche durchzuführen. Ziel von Datenbankanbietern sollte es daher sein, die typischen Recall-Werte von Recherchen in einer Datenbank durch erweiterte Suchszenarios mit mehreren Datenbanken zu erhöhen. Die Zusammenführung von Datenbanken über das Konzept der Heterogenitätsbehandlung, das in dieser Arbeit evaluiert wurde, kann als ein tragfähiges Konzept für die datenbankübergreifende Online-Recherche angesehen werden. Die Ergebnisse zeigen, dass sich der Recall beim Einsatz von Crosskonkordanzen automatisch erhöht, was bei einer Recherche im wissenschaftlichen Umfeld i. d. R. angestrebt wird. Ungeklärt ist, ob sich das Verfahren Crosskonkordanzen auch positiv auf informatrische Untersuchungen auswirkt. Prinzipiell ist unstrittig, dass sich eine Recall-Erhöhung auch für informatrische Untersuchungen und deren Aussagekraft positiv auswirken müsste (siehe in diesem Zusammenhang die Aussagen zur Situation der Forschungsrankings bei (Hornbostel, 2006)). Die Unvollständigkeit (vgl. geringer Recall) von Bibliographien, die beispielsweise häufig als Grund für Abweichungen (z. B. Groos droop) von informatrischen Gesetzmäßigkeiten genannt wird, ließe sich durch die Anwendung von Crosskonkordanzen abmindern. Auf das nicht zu vernachlässigende Problem der Dokumenten-Dubletten bei einer Recherche in mehreren fachlich relevanten Datenbanken soll hier nicht weiter eingegangen werden (siehe dazu (Hood, 1999, Hood und Wilson, 2003)).

Es muss erwähnt werden, dass bei informatrischen Untersuchungen im Allgemeinen nur die Publikationen mitberücksichtigt werden können, die sozusagen „visible“ sind, so wie im Internet-Bereich das Visible Web vom Invisible bzw. Deep Web unterschieden wird ((Bergman, 2001, Lewandowski und Mayr, 2006)). Dass beispielsweise wissenschaftliche Publikationen in China in der westlichen Welt kaum adäquat wahrgenommen werden, aber in immer größerem Umfang dort entstehen (Jin und Wang, 1999), darf bei diesen Überlegungen nicht vergessen werden.

Die Entwicklung bei den Zitationsdatenbanken – inzwischen existiert neben dem WoS mind. eine weitere internationale Zitationsdatenbank (Scopus⁸⁹) – zeigt, dass künftig u. U. auf mehrere Zitationsdatenbanken zugegriffen werden muss, um umfassende informatrische Untersuchungen durchführen zu können (siehe dazu (Bar-Ilan, 2006, Bar-Ilan, 2008)). Auf die Abdeckung der elektronischen Publikationen durch das Suchsystem Google Scholar wurde bereits hingewiesen (siehe Abschnitt 2.7).

89 <http://www.scopus.com>

Zusammenfassend lässt sich sagen, dass das hier evaluierte Konzept der Crosskordanzen, das generell für die verteilte Datenbankrecherche konzipiert wurde, auch für andere Einsatzszenarien passfähig ist. Die Unterstützung von informetrischen Untersuchungen soll hier als ein Beispiel genannt werden. Da informetrische Untersuchungen von Bibliographien ebenfalls auf deskriptoren-basierten Recherchen in Fachdatenbanken basieren (z. B. (Mayr und Umstätter, 2007, Peritz, 1990, Wilson, 1995)), ist die Nützlichkeit für diesen Typ Recherche sehr plausibel.

8.2 Bradford-Verteilungen für Zeitschriften- und Monographienliteratur

Inwieweit lassen sich Bradford-Verteilungen in unterschiedlichen Fachgebieten und insbesondere bei Recherchen in Literaturdatenbanken identifizieren?

Der folgende Abschnitt bezieht sich auf die Ergebnisse in Abschnitt 7.2.

Bradford-Verteilungen treten erfahrungsgemäß in allen Fachgebieten und damit auch bei Recherchen in Literaturdatenbanken auf. Diese These kann für den überwiegenden Teil der Fragestellung der beiden Evaluationsprojekte bejaht werden. Die Zeitschriftenartikel zu den IR-Topics aus CLEF und KoMoHe nähern sich der typischen Bradford-Verteilung an (siehe dazu Tabellen und Abbildungen in Abschnitt 7.2.1). Die zusätzlichen Tests in Kapitel 6 bestätigen dies für weitere in der Hauptuntersuchung nicht berücksichtigte Datenbanken und Operationalisierungen (siehe dazu auch (Hjørland und Nicolaisen, 2005, Nicolaisen und Hjørland, 2007)). Es wurden folgende Operationalisierungen und Datenbanken untersucht: zufällig ausgewählte Ein-Wort-Anfragen in den Datenbanken SOLIS, WAO und CSA-SA (siehe Abschnitt 6.2), breit operationalisierte Themen (sozialwissenschaftliche Fragestellungen) in den Datenbanken SOLIS, IBLK, WPSA, CSA-SA, PEI, ASSIA (siehe Abschnitt 6.3) sowie die Literatur zum Forschungsthema Informetrics in der Datenbank LISA (siehe Abschnitt 6.4).

In der Literatur finden sich viele weitere Beispiele für Studien, die Bradford-Verteilungen in unterschiedlichen Domänen und Fachgebieten nachweisen und damit auf die Generalisierbarkeit dieses Verteilungstyps schließen lassen (siehe z. B. (Cline, 1981)). Beispielsweise finden sich Untersuchungen von Bradford-Verteilungen in folgenden Bibliographien und Studien:

- Fragestellungen in sozialwissenschaftlichen Bibliographien bei (Coleman, 1993, Coleman, 1994),
- das Fachgebiet „Bibliometrics“ (Peritz, 1990),
- das Topic „Information retrieval and seeking“ in unterschiedlichen Datenbanken (von Ungern-Sternberg, 2000),
- Literatur zum Thema „Emergency management“ (Tenopir, 1982),
- das Topic „Classification Systems Literature“ in (Chung, 1994),
- in den Disziplinen Logik (20. Jahrhundert) und Mathematik (19. Jahrhundert) bei (Wagner-Döbler, 1997),
- das Thema „Methanogenic bacteria“ (Pontigo und Lancaster, 1986),
- das Fachgebiet der Geographie (Aiyepku, 1977),

- das Topic „Noble gas compounds“ (Hawkins, 1978),
- das Fachgebiet der „Oceanography“ (Tapaswi und Maheswarappa, 1999),
- das Fachgebiet der „Biomedical communications“ (Triolo und Bao, 1993) und
- das Fachgebiet der „Semiconductor Literature“ (Tsay et al., 2000).

In der Literatur finden sich viele weitere disziplinübergreifende Beispiele. Diese Studien beschreiben verschiedene Typen von Verteilungen (z. B. Bradford ähnliche, vgl. (Bookstein, 1990a, Bookstein, 1990b)), die für unterschiedliche Dokumenttypen (z. B. Zeitschriftenliteratur, Monographien, Datenbanken, Domain-Namen) gelten.

Die Mehrzahl der zuvor aufgezählten Studien untersucht ein spezifisches Thema in einer oder mehreren Datenbanken. Insbesondere die Fragestellungen in den Naturwissenschaften (z. B. bei (Hawkins, 1978, Pontigo und Lancaster, 1986, Triolo und Bao, 1993)) konzentrieren sich, wie ursprünglich auch bei Bradford (1934), auf sehr spezifische Forschungsthemen, die anhand einer umfassenden Bibliographie untersucht wurden. Die meisten Themen, die vom Autor in dieser Arbeit untersucht wurden, stammen thematisch aus den Sozialwissenschaften (insbesondere CLEF-Topics in Abschnitt 5.4.1). Die Fragestellungen der CLEF-Reihen und KoMoHe-Reihen (Test 1 und Test 2) zeigen, dass sich die Bradfordsche Gesetzmäßigkeit erfolgreich auch auf sozialwissenschaftliche Fragestellungen und verwandte Domänen übertragen lässt. Die Besonderheit dieser Untersuchung liegt darin, dass für eine große Anzahl von sehr unterschiedlichen Fragestellungen die Konzentration der Literatur in vergleichsweise kleinen (unvollständigen) Dokumentkorpora untersucht wurde. Die relativ geringen Dokumentzahlen der CLEF-Topics bei Zeitschriftenartikeln (siehe Tabelle 41) zeigen, dass auch in unvollständigen Bibliographien (der GIRT-Korpus enthält lediglich deutschsprachige Forschungsliteratur der Jahre 1990 - 2000) die Bradford-typischen Verteilungen auftreten. Es ist für beide Untersuchungsreihen (CLEF und KoMoHe) davon auszugehen, dass die untersuchten Bibliographien nur einen Teil der verfügbaren Forschungsliteratur beinhalten. Der GIRT-Korpus und die Datenbank-Ausschnitte der KoMoHe-Evaluation sind insofern unvollständig, als dass die Ergebnismengen für den IR-Test künstlich begrenzt werden (vgl. (Kluck, 2002, Voorhees, 2007)). Trotz dieser Begrenzung (Unvollständigkeit der Bibliographien) zeigen sich die für Bradford typischen Verteilungen.

Während die gemittelten Zahlen und Bradford-Graphen der fünf CLEF-Jahrgänge (siehe Tabelle 41) relativ homogen verlaufen, zeigen die KoMoHe-Topics in den drei unterschiedlichen Testreihen deutlichere Unterschiede bei der Größe der drei Bradford-Zonen und dem Verlauf der Graphen. Die Unterschiede zwischen den KoMoHe-Testreihen lassen sich vor allem auf die Dokumentmenge zu den einzelnen Topics und die unterschiedlich großen Datenbanken zurückführen. Die Dokumentzahlen in der Datenbank Medline und Psyn dex (Test 3, siehe Tabelle 42) liegen im Mittel knapp doppelt so hoch wie die Zahlen in den übrigen Datenbanken im KoMoHe-Test (Test 1 und 2). Dies hat Auswirkungen auf die gemittelte Anzahl der Zeitschriften im Core und den nachfolgenden Zonen in dieser Testreihe. Von einer starken Konzentration der Zeitschriftenliteratur in den naturwissenschaftlichen Fächern auf wenige Kernzeitschriften kann auf Basis der Analyse der Topics in Test 3 nicht gesprochen werden. Die 12 Topics in Test 3 ergeben im Mittel 16 Kernzeitschriften, während die Topics in Test 1 und Test 2 nur rund 8 bzw.

6 Kernzeitschriften ergeben (siehe Tabelle 42), was für eine stärkere Konzentration in Test 1 und Test 2 sprechen würde. Zum Teil lässt sich dieses Ergebnis auch dadurch erklären, dass insbesondere in der Medline-Datenbank viele tausend Zeitschriften erfasst werden, während in der kleineren sozialwissenschaftlichen Datenbank SOLIS, die Basis des GIRT-Korpus ist, nur rund 300 Zeitschriften ausgewertet werden. Ein Topic in einer sehr großen Datenbank wie Medline kann also theoretisch deutlich stärker streuen als ein Thema in einer vergleichsweise kleinen Datenbank, die nur wenige Zeitschriften erschließt. Die stärkere Streuung eines Themas in einer Datenbank kann hierbei mehrere Gründe haben. Z. B. hat die Erschließungstiefe und -politik der Datenbank großen Einfluss auf die Streuung eines Themas (siehe dazu Untersuchung der Datenbank LISA in (Mayr und Umstätter, 2007)). Zum anderen kann die Anfrageformulierung bei der Dokumentengewinnung zu allgemein oder auch zu spezifisch operationalisiert sein und somit die Dokumentenmenge beeinflussen. So konnten aus den insgesamt 82 Topics der KoMoHe-Evaluation (siehe Abschnitt 5.4.2) für die Bradfordizing-Analyse beispielsweise nur 39 Topics verwendet werden, weil diese 39 Dokumentenpools die notwendige Anzahl an Dokumenten ergaben. Die übrigen Topics waren zu spezifisch formuliert oder außerhalb des Scopes der analysierten Datenbank und ergaben unter 100 Dokumente, was für ein Bradfordizing i. d. R. nicht ausreicht. Der vergleichsweise hohe Wert der Testreihe 3 in der KoMoHe-Evaluation erklärt sich zusätzlich durch zwei „Ausreißer-Topics“, die den Mittelwert etwas zu hoch erscheinen lassen. Die beiden Topics *psynindex-mesh3* und *psynindex-mesh5* (bzw. Topic-Nummer 363 und 365 in Tabelle 23) **ergeben Zeitschriftenartikel** in 822 bzw. 643 Zeitschriften (vgl. Abbildung 7-8), was verglichen mit den übrigen Topics im Test ein deutlich zu hoher Wert ist. Tabelle 42 spiegelt also weniger Konzentrationsunterschiede zwischen den Wissenschaftsdisziplinen wider, sondern vielmehr Spezifika/Artefakte bei der Anfrageformulierung und der Datenbankauswahl.

Inwiefern die Operationalisierung Einfluss auf die Belegung der Bradford-Zonen nehmen kann, zeigt sich anhand von exemplarischen Tests bei (Nicolaisen und Hjørland, 2007). Nicolaisen & Hjørland haben Topics auf unterschiedliche Arten für eine Datenbankrecherche operationalisiert und erhalten nach der Analyse der Dokumentmenge der jeweiligen Operationalisierung unterschiedliche Rangfolgen bei den Zeitschriften.

Lassen sich diese für Zeitschriftenliteratur typischen Verteilungen auch bei anderen Dokumenttypen (z.B. Monographien) beobachten?

Zusätzlich zu den Zeitschriften wurde ein Teil der CLEF- und KoMoHe-Topics auch für die Monographienliteratur untersucht. Die bei (Brookes, 1969) und später bei (Worthen, 1975) beschriebene Übertragbarkeit der Bradfordschen Gesetzmäßigkeit auf die Monographienliteratur kann durch diese Studie für die analysierten Topics auf einer sehr breiten Basis bestätigt werden (siehe Tabelle 43 und Tabelle 44). Es zeigt sich, dass sich das in dieser Arbeit entwickelte Vorgehen der Analyse der Verlags-Codes aus der ISBN der Monographie-Datensätze als tragbares Verfahren erweist.

Tabelle 43 verdeutlicht, dass die Monographienliteratur rein mengenmäßig in den Sozialwissenschaften eine wichtige Rolle spielt. Der Vergleich der gemittelten Werte der drei Testreihen mit den Werten für die Zeitschriftenliteratur (siehe Tabelle 41) ergibt

einen etwas höheren Anteil an Monographien bei den CLEF-Topics (vgl. dazu (Dees, 2008) für die Forschungsliteratur in der Pädagogik). Die Testreihe 3 im KoMoHe-Test (Test 3, siehe Tabelle 44) zeigt, dass der Anteil der Monographien an der Forschungsliteratur in den nicht-sozialwissenschaftlichen Disziplinen (insbesondere Medizin und Psychologie in Test 3) deutlich geringer ist. In Medline werden überwiegend Zeitschriftenartikel ausgewertet, was sich in Test 3 deutlich zeigt.

Auffällig bei der Analyse der Monographienliteratur ist weiterhin, dass die Größe der drei Bradford-Zonen sehr ähnlich verläuft wie bei der Zeitschriftenliteratur. Unabhängig von der Dokumentenart (Zeitschriftenartikel oder Verlagsmonographien) finden sich etwa 6 - 7 Zeitschriften oder Verlage, die den Kern der Literatur für ein Topic zur Verfügung stellen. Für die zweite Zone werden 3- bis 4-mal so viele Zeitschriften oder Verlage benötigt. Für die dritte Zone werden zwischen 2- bis 3-mal so viele Zeitschriften oder Verlage benötigt, um das letzte Drittel der Literatur bereitzustellen.

Dieses Ergebnis verdeutlicht die Generalisierbarkeit dieses Verteilungstyps für unterschiedliche Dokumentkorpora, Dokumenttypen und Fachdisziplinen auf der Basis von 164 Fragestellungen. Abbildung 7-15 zeigt sehr deutlich, dass die Steigungsexponenten von 100 Graphen aus den Testreihen (CLEF und KoMoHe) unabhängig vom Dokumenttyp und der Topic-Größe bei etwa 0,5 liegen und damit einer Power Law-Verteilung folgen.

Weitere Spezifika der unterschiedlichen Streuung zwischen den einzelnen Topics und tiefergehende informatrische Untersuchungen, wie z.B. die bei (Wagner-Döbler, 1996, Wagner-Döbler, 1997) untersuchten Bradford-Verteilungen in unterschiedlichen Zeitspannen und Entwicklungsständen einer Disziplin (siehe dazu (Sen und Chatterjee, 1998)), konnten in dieser Arbeit aufgrund der Anzahl der analysierten Topics nicht durchgeführt werden.

8.3 Mehrwert durch Bradfordizing

Lassen sich informatrische Verfahren (z. B. Bradfordizing) als Mehrwertdienste für die Informationssuche in Literaturdatenbanken gewinnbringend nutzen?

Der folgende Abschnitt bezieht sich auf die Ergebnisse in Abschnitt 7.3.

Da Fragen nach der Vertrautheit, Kenntnis und Relevanzeinschätzung von Zeitschriften zu einem Thema idealiter anhand von konkreten Fragestellungen untersucht werden, haben sich die CLEF-Topics aus dem letzten Jahr (2007) als Befragungsgegenstand angeboten. Insgesamt konnten 24 Fragestellungen (Topics Nr. 177 - 200; siehe Tabelle 21 in Abschnitt 5.4.1) mit Personen der GESIS in Verbindung gebracht werden. Das Kurzinterview, das keinen Anspruch auf ein repräsentatives Ergebnis erhebt, kann einige interessante Hinweise für die Diskussion liefern.

Die Antworten zu Frage 1 „spontane Nennung von Kernzeitschriften“ zeigen, dass es selbst für erfahrene Informationsspezialisten schwierig sein kann, spontan die Kernzeitschriften zu einem bekannten Thema aus einem Pool von ca. 300 möglichen Zeitschriften zu benennen (siehe Tabelle 45). Das Bradfordizing einer thematischen Treffermenge kann damit auch für erfahrene Informationsfachleute einen informationellen Mehrwert generieren, zumal sich die Kernzeitschriften, die rein quantitativ durch Bradfordizing

bestimmt wurden, u. U. von favorisierten Zeitschriften (z. B. hohe Reputation) unterscheiden können.

Der Erkenntnisgewinn selbst für den Nutzertyp „Informationsspezialist“ kann durch Frage 2 verdeutlicht werden (siehe Tabelle 46). Verallgemeinernd lässt sich eine deutliche Diskrepanz zwischen sehr erfahrenen „Experten“ und eher „laienhaften“ Bearbeitern unter den Befragten feststellen. Für ungeübte Nutzer bzw. Nutzer, die sich einen Überblick über eine Dokumentenmenge verschaffen wollen, kann das Verfahren Bradfordizing vor allem eine zusätzliche Strukturierung einer Treffermenge ermöglichen. Ein nach Zeitschriften strukturiertes Ergebnis unterstützt zudem bei der Navigation der Dokumente (vgl. (Bates, 1989, Bates, 2002)). Es ist anzunehmen, dass vermutlich Anfänger und Fachfremde den größten Mehrwert durch das Verfahren Bradfordizing erhalten. Mehrere Studien des Forschungsgebiets Informetrics (siehe z. B. (Mayr und Umstätter, 2007)) haben jedenfalls für aktiv Forschende in den Informationswissenschaften keine unbekannten Zeitschriften im Kern gezeigt; u. U. ist daher für diesen Personenkreis insbesondere die Zone 3, der sog. long tail der Verteilung, interessant. Auch für diesen Fall kann das Bradfordizing einen Mehrwert generieren (z. B. Listen aller Zeitschriften und Artikel in Zone 3).

Frage 3 „Relevanz der Kernzeitschriften“ fokussiert auf die Einschätzung der Relevanz der Kernzeitschriften gegenüber den übrigen Zeitschriften (siehe Tabelle 47). Eine Mehrheit der Befragten ist intuitiv von dem Relevanzvorteil der Kernzeitschriften überzeugt. Es zeigt sich weiterhin eine Diskrepanz zwischen zuerkannter Relevanz der Kernzeitschriften und ihrer aktiven Nutzung. Dies kann ebenfalls an der Art und Thematik der wenigen Topics oder der Fragesituation liegen. Es fällt auf, dass für spezifischere Topics (z. B. Umwelt-Soziologie, Gender- und Frauenforschung usw.) leichter Kernzeitschriften zu benennen waren als für allgemeine, sozialwissenschaftliche Themen wie z. B. Integrationspolitik.

Weitere Hinweise der Interviewpartner lassen sich folgendermaßen zusammenfassen:

- Einige Interviewpartner gaben bei Frage 4 Zeitschriften an, die erst in den letzten Jahren gegründet bzw. in die Datenbank aufgenommen wurden und daher nicht in den analysierten Zeitschriften vorkommen konnten (Zeitraum 1990 - 2000). Diese „neuen“ Kernzeitschriften seien hochrelevant zu der Thematik, aber nicht im Sample (siehe dazu (Brookes, 1971)). Z. T. wurden auch nicht-sozialwissenschaftliche Zeitschriften genannt (Bezug zu These 1). Es wurden zudem einige Zeitschriften genannt, die in der zugrunde liegenden Datenbank SOLIS nicht erfasst werden (z. B. statistische Zeitschriften, die nur gelegentlich etwas zum Thema beinhalten).
- Eine Person gibt an, dass für sie Personen, also Autoren, relevantere „thematische“ Identifikatoren sind als Kernzeitschriften.
- Zeitschriften werden von einer anderen Person gerne auch als Ausschlusskriterium bei der Dokumenten-Recherche verwendet.
- Eine Person, die keine Kernzeitschriften benennen kann (Frage 1), gibt nach dem Vorlesen der Kernzeitschriften an, dass die Titel der Zeitschriften alle sehr plausibel und relevant klingen.

- Einige Topics waren sehr interdisziplinär und eine Person äußert sich, deshalb überfordert zu sein, konkrete Kernzeitschriften zu nennen.

Eine wöchentlich erscheinende Zeitschrift „Aus Politik und Zeitgeschichte“ – die Beilage zur Wochenzeitung „Das Parlament“ (APuZ) erscheint häufiger in der Liste der Kernzeitschriften bei einigen Topics. Es wird von einer Person auf die häufige Erscheinungsweise hingewiesen; es wird aber auch erwähnt, das APuZ als Kernzeitschrift für mehrere Themen gelten kann.

Die Antworten zu Frage 4 nach den weiteren fachlich relevanten Zeitschriften unterstützen das in diesem Abschnitt angesprochene Problem der unvollständigen Bibliographien (siehe auch Groos droop in Kapitel 4). Zu knapp der Hälfte der Topics können weitere Zeitschriften genannt werden, die entweder nicht in der Datenbank erschlossen werden (z. B. außerhalb des Scopes) oder erst in den letzten Jahren gegründet wurden und daher nicht im Sample enthalten waren.

Ein Relevanzvorteil der Kernzeitschriften wird im Allgemeinen deutlich bejaht (vgl. These 6 in Kapitel 10). Bemerkenswert ist, dass alle Personen, die Kernzeitschriften kennen, auch von der Relevanz dieser Zeitschriften überzeugt sind.

8.4 Relevanzvorteil durch Bradfordizing bei Zeitschriftenartikeln und Monographien

Verbessert das Bradfordizing den Anteil der relevanten Dokumente im ersten Drittel (Core) einer Dokumentmenge gegenüber dem zweiten und dritten Drittel (Zone 2 bzw. Zone 3)?

Der folgende Abschnitt bezieht sich auf die Ergebnisse in Abschnitt 7.4.

Die Auswertung der Relevanzurteile für den Dokumenttyp Zeitschriftenartikel zu den 125 Topics der CLEF-Reihe zeigt ein klares Ergebnis. Die Dokumente, die nach Bradfordizing in die Kernzone (core) gerankt werden, sind im Mittel häufiger relevant für die Fragestellungen als Dokumente, die sich in Zone 2 oder Zone 3 befinden. Die Visualisierungen der Precision-Werte pro Zone und Topic in den Abbildungen (Abbildung 7-16 bis Abbildung 7-20) verdeutlichen den Precision-Vorteil der Dokumente aus den Kernzeitschriften für den Großteil der Topics. Einige Topics zeigen eine andere Relevanz-Verteilung. Zu diesen Topics finden sich beispielsweise mehr relevante Dokumente in Zone 2 oder Zone 3. Die gemittelten Verbesserungen in Tabelle 50 zeigen Verbesserung zwischen dem Core und Zone 3, Core und Zone 2, Core und baseline (Durchschnitt) sowie zwischen Zone 2 und Zone 3. Für die meisten Testreihen können signifikante Verbesserungen nachgewiesen werden (siehe *-Symbol in Tabelle 50). Vergleichsweise geringe Verbesserungen auf nichtsignifikantem Niveau zeigt die Testreihe aus dem Jahr 2006.

Der Relevanzvorteil des ersten Drittels (core) eines Suchergebnisses, auf den es den meisten Nutzern vermutlich ankommen wird, kann durch die CLEF-Reihe aufgezeigt werden (vgl. dazu (Brookes, 1971)). Ein Nutzer, der sich vorrangig mit den Dokumenten im Core beschäftigt, kann davon ausgehen, dass er im Core ca. 56 % mehr relevante Dokumente zu seiner Anfrage finden wird als im letzten Drittel der Ergebnismenge. Die Verbesserung des Core zu einem zufällig sortierten Gesamtergebnis (alle bewerteten

Dokumente einschließlich Core, siehe baseline-Werte) liegt immerhin bei 22 %. Interessant wäre an dieser Stelle ein Vergleich der Precision-Verteilung für boolesche und best-match Ranking-Verfahren gegenüber der Bradfordizing-Methode. Dieser Vergleich konnte in dieser Arbeit nicht durchgeführt werden und soll daher zu einem späteren Zeitpunkt erfolgen.

Die Ergebnisse in der KoMoHe-Reihe bestätigen das Ergebnis der CLEF-Fragestellungen tendenziell, allerdings kann der Relevanzvorteil im Core nur auf einem etwas niedrigeren Niveau verifiziert werden. Insgesamt können niedrigere Verbesserungsraten nachgewiesen werden (siehe Tabelle 52). **Insbesondere die Topics in Test 3 zeigen signifikante Verbesserungen zwischen den Zonen.** Für die Topics in Test 1 und Test 2 liegen die Verbesserungen prozentual auf einem ähnlichen Niveau wie Test 3, aufgrund einiger Abweichungen können aber für Test 1 und Test 2 keine signifikanten Verbesserungen präsentiert werden.

Lässt sich das Verfahren Bradfordizing als Re-Ranking-Ansatz auch erfolgreich auf andere Dokumenttypen übertragen?

Der folgende Abschnitt bezieht sich auf die Ergebnisse in Abschnitt 0.

Die Auswertung der Relevanzurteile für den Dokumenttyp Monographien zu den 75 Topics der CLEF-Reihe und 30 Topics der KoMoHe-Reihe zeigt ein ähnliches Ergebnis wie die Auswertungen der Zeitschriftenartikel. Die Verbesserung der Precision nach der Anwendung des Re-Rankings auf Basis von Bradfordizing kann bestätigt werden. Allerdings fallen die Verbesserungen der CLEF-Reihe niedriger aus als beim Dokumenttyp Zeitschriftenartikel. Es können keine signifikanten Verbesserungen bei den Monographien nachgewiesen werden. Die Auswertung der Relevanzurteile für Monographien in den drei Zonen im Datensatz KoMoHe zeigt signifikante Verbesserungen zwischen core und Zone 3 sowie zwischen Core zu baseline (siehe Tabelle 54).

Auffällig ist der minimale Anstieg der Precision für Monographien in Zone 3 bei beiden Testreihen (CLEF und KoMoHe), der für Zeitschriftenartikel nicht beobachtet werden konnte. Dieser Anstieg erklärt sich dadurch, dass Monographien-Nachweise ohne ISBN im Datensatz automatisch in die Zone 3 verschoben wurden, was zu einer vergleichsweise großen Zone 3 geführt hat. Als Folge dieses Verfahrens gleicht sich die Precision der Zone 3 der baseline-Precision an. Es kann daraus nicht geschlossen werden, dass sich der Anteil der relevanten Dokumente in Zone 3 wieder erhöht.

Die Einteilung der Dokumente pro Topic in drei Bereiche oder Zonen ist eine beliebige Einteilung des Autors, die dem ursprünglichen Vorschlag von (Bradford, 1934) folgt. Möglicherweise ist die Zoneneinteilung in mehr als drei Zonen, wie bei (Egghe und Rousseau, 1990) vorgeschlagen, für die Ergebnisse des Re-Ranking hilfreich. Der Nutzer hätte bei diesem Vorgehen (z. B. 7 Zonen) deutlich weniger Dokumente im Core; es ist aber fraglich, ob sich eine Verkleinerung des Core auch positiv auf die Precision der Zonen auswirkt.

Eine Ausweitung der Bradfordizing-Analyse auf weitere Dokumenttypen erscheint operabel. Beispielsweise kann das Verfahren Bradfordizing u. U. auch auf Sammelwerke

(anhand der ISBN), Sammelwerksbeiträge oder Forschungsprojektnachweise⁹⁰ angewendet werden. Die Bibliogram-Analyse bei (White, 2005) für seed-terms oder Deskriptoren und die Weiterentwicklung dieses Verfahrens in Form von Pennant-Diagrammen⁹¹ (White, 2007a, White, 2007b) erscheint in diesem Zusammenhang auch als lohnend.

In (Mayr et al., 2008) wurde zudem eine Koppelung von Bradfordizing mit einem weiteren Verfahren der Scientometrie vorgeschlagen (Re-Ranking nach Autorenzentralität⁹²). Ziel des Dokumenten-Ranking nach Autorenzentralität ist es, Veröffentlichungen zentraler Autoren zu finden (zentrale Dokumente). Die Grundannahme dieses Modells ist, dass die Relevanz von Publikationen mit der Zentralität ihrer Autoren steigt. Das Ranking bei (Mutschke, 2004) geht davon aus, dass die Wahrscheinlichkeit, (mehr) relevante Dokumente zu finden, innerhalb der Gruppe der zentralen Autoren größer ist als innerhalb der Gruppe der Autoren, die eher an der Peripherie eines Autorennetzwerkes angesiedelt sind (was aber nicht bedeutet, dass Publikationen nicht-zentraler Autoren nicht relevant sind). Eine Kombination oder Koppelung der beiden Verfahren findet sich in Abbildung 8-3.

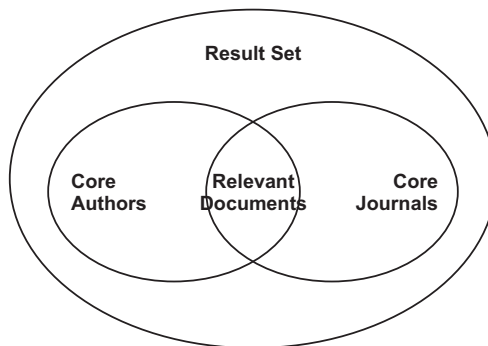


Abbildung 8-3: Koppelung von Kernzeitschriften und zentralen Autoren (aus (Mayr et al., 2008))

Die Evaluation der Retrieval-Effektivität von Bradfordizing auf empirischer Basis (siehe Abschnitt 7.4.1) stellt bis dato ein Forschungsdesiderat dar. In den aktuellen Arbeiten zur

90 Da der Dokumenttyp Forschungsprojekt für die Bradfordizing-Analyse nicht unmittelbar zur Verfügung steht, kann über die Autoren (Projektleiter und Mitarbeiter) des Projekts und deren Veröffentlichungen die Analyse durchgeführt werden.

91 „Pennant diagrams use bibliometric data and IR techniques on the system side to mimic a relevance-theoretic model of cognition on the user side.“

92 Grundlage hierfür sind soziale Netzwerke, die sich durch Kooperation wissenschaftlicher Autoren konstituieren (Autorennetzwerke, vgl. Beaver, D. deB. (2004): Does collaborative research have greater epistemic authority? *Scientometrics* 60 [3], Seite 399-408.). Hauptanwendungsfeld dieser Autorennetzwerkmodelle ist die Suche nach Experten und das Ranking von Dokumenten auf der Basis von Autorenzentralität. Zentralität ist dabei ein strukturelles Merkmal, das den Beitrag der Netzwerkposition eines Akteurs zu dessen Prominenz, Einfluss oder Wichtigkeit in der sozialen Struktur misst. Als Zentralitätsmaße wurden degree, closeness und betweenness verwendet (vgl. Wasserman und Faust, 1994).

Relevanztheorie (Saracevic, 1975, Saracevic, 2007a, Saracevic, 2007b, Sperber und Wilson, 1995, White, 2007a, White, 2007b) wird explizit auf die Bedeutung von Bradford's Law of Scattering für die Relevanzforschung eingegangen. Eine empirische Untersuchung der Auswirkungen dieser Verteilung anhand von konkreten Fragestellungen auf Retrieval-Ergebnisse bleibt aber aus. Diese Lücke kann durch die Ergebnisse in dieser Arbeit geschlossen werden.

Definitorische Charakterisierungen von Relevanz finden sich bei unterschiedlichen Autoren. (Mizzaro, 1997) beschreibt beispielsweise sieben Aspekte von Relevanz. Einzelne ausgewählte Zitate sollen das Konzept Relevanz, das in den Laborexperimenten dieser Arbeit eine große Rolle spielt, illustrieren:

“Relevance is considered as a measure of the effectiveness of a contact between a source and a destination in a communication process. (...) In the most fundamental sense, relevance has to do with effectiveness of communication.” (Saracevic, 1975) S. 321

“Relevance is a property which need not be represented, let alone computed, in order to be achieved. When it is represented, it is represented in terms of comparative judgements and gross absolute judgements, (e.g. ‘irrelevant’, ‘weakly relevant’, ‘very relevant’), but not in terms of fine absolute judgements, i.e. quantitative ones.” ((Sperber und Wilson, 1995) S. 132)

“... relevance is intangible ... relevance is understood intuitively ... relevance is tacit” ((Saracevic, 2007) S. 1916)

“... relevance is one of the central concepts, if not the central concept, for documentation, information science, and information retrieval ... Relevance is not a well understood concept ...” ((Mizzaro, 1997) S. 810)

Saracevic sieht Relevanz sogar als den Schlüsselbegriff, der zur Entstehung der Informationswissenschaft und den informationswissenschaftlichen Theorien (z.B. Bradford's Law of Scattering) geführt hat ((Saracevic, 1975) S. 323). Er erklärt seine These mit dem zahlenmäßig starken Zuwachs an Wissenschaftlern und deren Publikations-Output (vgl. dazu (Nourmohammadi und Umstätter, 2004, Umstätter, 2004)), mit der daraus folgenden Notwendigkeit des Einzelnen, eine Auswahl zu treffen, mit dem Verschwimmen von Fächergrenzen (Interdisziplinarität) und mit dem zugleich steigenden Spezialisierungsgrad in den Wissenschaften. Er sieht Bradford als einen der Begründer der modernen Dokumentation, besonders weil Bradford als erster den Relevanzbegriff in einem informationswissenschaftlichen Sinne thematisiert („relevant to a subject“ bei (Bradford, 1934)) und anwendet hat.

„As far as I can determine, S.C. Bradford was the first one to use the term relevant in the context that it is used today in information science.” ((Saracevic, 1975) S. 324)

Bibliometrische Verteilungen bezeichnet Saracevic folglich als „relevance related distributions“. Diese Aussage trifft damit den Kern der Untersuchung in Abschnitt 7.4.1, in dem die Relevanzverteilungen zu fachlichen Fragestellungen untersucht werden.

Die folgende Aussage zu den unterschiedlichen Aspekten der Relevanz, die sich später auch bei (Mizzaro, 1997) finden, tangieren die Auswertungen und Messungen zur Relevanz in dieser Arbeit:

“... relevance should be considered independently ... relevance is multivalued ... not a simple yes/no decision ... Relevance of given documents may change as a result of other documents ...” (1975: S. 327)

Das in Abschnitt 5.2 angewendete „klassische“ IR-Laborexperiment mit binären Relevanzurteilen und den resultierenden Standardmaßen Recall und Precision ist nach dieser Aussage angreifbar (vgl. dazu auch (Chu, 2003, Ingwersen und Järvelin, 2005)). Die Juroren, die die Relevanzbewertung durchführen, sind dazu gezwungen, binäre Urteile zu treffen und ändern ihre Urteile i. d. R. nach einem Bewertungsdurchgang nicht mehr. Auch wurden die Dokumente in den Experimenten nur von einer Person bewertet. Ein erneutes Bewerten u. U. von anderen Personen ist aus zeittechnischen Gründen unüblich. Weitere Faktoren, die die Relevanzurteile beeinflussen können, finden sich auch in unseren Experimenten: unterschiedliche Repräsentation der Dokumente, unterschiedliche Spezifität der Fragestellungen, unterschiedliche Interpretation der Bewertungssituation und Ausdrucksmöglichkeiten bei den Bearbeitern, die die Bewertung durchführen. Es kommt in unserem Ansatz hinzu, dass Benutzeranfragen nur bedingt über kontrollierte Vokabulare operationalisiert werden können (klassisches Information Retrieval-Problem), daher sind alle Verfahren, die auf kontrollierten Vokabularen basieren oder diese erweitern, mit einer gewissen Unschärfe und Vagheit versehen. Trotz dieser Einschränkungen beim Bewerten von Dokumenten findet sich bei Saracevic folgende Einschätzung von Relevanzurteilen:

“One of the major conclusions that can be drawn from experiments is that relevance judgments are not at all associated with a random distribution. Although it may appear that relevance judgment is a very subjective human process, it has associated with it some remarkable regularity patterns.” ((Saracevic, 1975) S. 342)

Eine ausführliche Diskussion der gegensätzlichen Sichtweisen oder Modelle des Information Retrieval findet sich bei (Ingwersen und Järvelin, 2005). Die Autoren vergleichen systemzentriertes (klassische Laborexperimente) vs. nutzerzentriertes IR und arbeiten Vor- und Nachteile der beiden Ansätze heraus. Im Anschluss präsentieren sie einen neuen integrativen Ansatz, der beide Relevanzmodelle (user relevance und systems relevance) vereinbart. Dieses bei Ingwersen & Järvelin vorgeschlagene Modell stellt einen sehr ambitionierten Ansatz dar, der selbst von den internationalen Retrieval-Konferenzen noch nicht internalisiert und adaptiert wurde. Es ist auch fraglich, ob das von den Autoren vorgeschlagene Framework in Gänze operationalisierbar ist.

Da es ein Ziel von Retrieval-Systemen ist, neben den relevanten Dokumenten auch möglichst viele nützliche Dokumente in den wahrnehmbaren Dokumentenraum der Nutzer zu bringen, scheinen die bisherigen „einfachen“ Relevanzmodelle nicht systematisch zu versagen. Ein Dokument, das zwar nur die beschreibenden Metadaten eines Fachdatenbanknachweises enthält, kann i. d. R. von fachkundigen Nutzern intuitiv zu einer objektiven Fragestellung bewertet werden. Dass Nutzer auch mit weniger und unstrukturierter Information Relevanzurteile treffen können, zeigen die Untersuchungen von Suchmaschinenergebnissen im Web IR (siehe Übersicht in (Henzinger, 2007, Lewandowski, 2005)). Es soll aber an dieser Stelle nicht negiert werden, dass Nachweise von Doku-

menten, die nach Autopsie der Nachweisinformation hochrelevant erschienen, sich bei der Lektüre des Volltextes als gänzlich irrelevant für die Fragestellung erwiesen haben.

Bei der Betrachtung der Relevanz-Konzentration in unseren Testreihen (siehe Tabelle 53 und Tabelle 55) fällt auf, dass es sich dabei nicht um eine massive Verschiebung der Relevanz hin zum Core handelt, sondern vielmehr um ein kontinuierliches Abfallen der durchschnittlichen Precision ausgehend vom Core. Der Core beinhaltet demnach signifikant häufiger relevante Dokumente als die nachfolgenden Zonen. Die Ergebnisse dieser Arbeit verifizieren und konkretisieren damit die Aussage von Saracevic (1975), der Bradford-Verteilungen als „relevance related distributions“ bezeichnet. Bradford selber, der zwar „relevant to a subject“ für seine Bibliographien postuliert, trifft explizit keine Aussage, die eine Relevanzkonzentration im Core nahelegen würde. Für Bradford sind im Prinzip alle Dokumente in den beobachteten Dokumentzonen relevant. Die Interpretation von Bradford's Law bei M. Bates geht in die gleiche Richtung:

“The number of content-relevant articles per zone is the same across all zones. The number of content-relevant articles per journal decreases from the nucleus outward. (...) The number of utility-relevant articles per zone increases from the nucleus outward (due to the impact of document familiarity). The number of utility-relevant articles per journal is (probably) flat across all zones.” ((Bates, 1996) S. 705)

Nach Kenntnis des Autors existieren vergleichsweise wenige Untersuchungen, die sich zu den qualitativen Aspekten der Bradford-Verteilung äußern (vgl. (Bookstein, 1994, Hjørland und Nicolaisen, 2005)). Die Studie von Pontigo & Lancaster stellt eine Ausnahme dar:

“When journals are ranked by decreasing numbers of articles published on some subject, all one can reasonably claim is that the journals lower down the list are less likely to produce articles on this topic than those in the higher positions. This does not necessarily imply that the less productive journals are of lower quality or that the articles they publish are less significant contributions to the subject under consideration. Indeed, to conclude this would be illogical because it would imply that specialized journals are inherently of higher quality than those more general in scope. To use but one example, the New England Journal of Medicine may contribute infrequently to the literature on some tropical disease but these rare contributions may be the most important items published on the subject.” ((Pontigo und Lancaster, 1986) S. 59-60)

Die Studie von Pontigo und Lancaster, die auf zwei unterschiedlichen unabhängigen Qualitätsmaßen (Expertenurteile und Zitationsanalysen) beruht, ergibt folgendes Ergebnis:

“The results suggest that no qualitative differences may exist among articles drawn from the various zones of a Bradford distribution but that differences among the zones may exist in terms of level of citation. The „fugitivity“ of items in the final zone reduces the chance that they will be found and, therefore, cited. Once found, however, they are as likely to be judged of high quality as items from the earlier zones. In this study, at least, level of citation is not a good predictor of „quality“ as determined by subject experts.” (1986: S. 70)

Die Fachdatenbanken z.B. WoS oder SOLIS, die die Grundsätze der Bradfordschen Gesetzmäßigkeit, also die Kernbereichsbildung in sich tragen und perpetuieren ((Garfield, 1996)), sind ein weiteres Beispiel für die Allgemeingültigkeit dieser „relevance related distributions“. Die Auswertung einer begrenzten Zahl an Zeitschriften und anderen Literaturtypen bezieht sich dabei notwendigerweise auf ein Fachgebiet und basiert auf einer intellektuellen Relevanzbewertung. Quellen, die außerhalb dieses Fachgebiets oder Scopes der Datenbank liegen, geraten automatisch in die Peripherie dieser Datenbank und sind aber nicht gänzlich irrelevant. Wäre die Peripherie für den Nutzer einer Datenbank irrelevant, würden die Verfahren der Heterogenitätsbehandlung keinen Mehrwert generieren, sondern Recall und Precision reduzieren. Gerade das Gegenteil ist bei der Untersuchung dieser Maße in Abschnitt 7.1 der Fall.

Zum Schluss der Diskussion soll auf eine aktuelle Untersuchung der Effekte von Bradfordizing eingegangen werden. Die Kritik an den Auswirkungen von Bradfordizing wird bei (Nicolaisen und Hjørland, 2007) folgendermaßen beschrieben:

„Bradford analyses function discriminatively against minority views ... Bradford analysis can no longer be regarded as an objective and neutral method.“

Die Autoren verkennen bei dieser verkürzten und empirisch nicht belegbaren These (es wurden lediglich 2 Topics analysiert), dass nicht die Bradfordizing-Analyse zur Diskriminierung von Meinungen und Publikationen neigt, sondern die Politik der Datenbankproduzenten bei der Erstellung von Datenbanken. Die Entscheidung eine Zeitschrift cover-to-cover auszuwerten oder sehr selektiv, liegt bei den Datenbankproduzenten; die Bradfordizing-Analyse, als rein quantitatives Verfahren, gibt letztlich nur den Stand der Datenbank wider. Richtig an der These ist aber, dass sehr „kleine“ Zeitschriften, die z. B. nur 10 Aufsätze im Jahr publizieren, eine viel geringere Chance haben, in den Bereich der Kernzeitschriften zu kommen.

Die empirischen Ergebnisse dieser Arbeit, die für eine Vielzahl von Fragestellungen eine Relevanzkonzentration im Core nachweisen, können damit als eine Konkretisierung der „relevance related distributions“ bei Bradford interpretiert werden. Bradfordizing ist somit in der Lage, die Relevanzhierarchie eines fachlichen Suchergebnisses für Literaturrecherchen signifikant zu verbessern. Letztlich existiert aber noch kein verifizierbares Erklärungsmodell für den Relevanzvorteil der Dokumente im Core einer Bradford-Verteilung. Die Aspekte von Relevanz sind dafür zu vielschichtig und wenig erforscht. Vermutlich spielen unterschiedliche Faktoren bei der Relevanzbewertung zusammen.

Plausibel ist, dass sich die Relevanz (gemessen in Precision) verbessert, weil Dokumente aus Kernzeitschriften spezifische Eigenschaften besitzen. So lässt sich folgendes über Kernzeitschriften aussagen:

- Kernzeitschriften repräsentieren i. d. R. den State of the Art eines Themas,
- Aufsätze in Kernzeitschriften durchlaufen i. d. R. das sog. peer-reviewing und werden von Spezialisten dieses Faches bewertet,

- Publikationen im Kernbereich decken mehrere Aspekte des jeweiligen Forschungsbereiches ab (siehe dazu Aussage bei (Bonitz, 1985))⁹³.

Das folgende Kapitel zeigt anschließende Forschungsfragestellungen auf. Dabei wird besonderer Wert auf die Kombination unterschiedlicher Mehrwertdienste gelegt.

93 “... Ranking by number of hits, i.e. the commonly used method, prefers multidisciplinary journals ...” (Bonitz, 1985: S. 482)

9 Ausblick

In der Arbeit konnten einzelne Forschungsfragen eingehender untersucht werden. Dabei hat sich gezeigt, dass die beiden Verfahren Crosskonkordanzen und Bradfordizing erfolgreich auf unterschiedliche Domänen und Dokumenttypen angewendet werden können. Die Effekte der einzelnen Verfahren konnten durch empirische Tests auf Basis einer klassischen IR-Evaluation positiv untersucht werden.

Das folgende Kapitel beschreibt darauf aufbauend ein Forschungsprojekt, das unterschiedliche Mehrwertverfahren (z. B. Bradfordizing) eingehender untersuchen wird und die Anschlussfähigkeit sowie Koppelung der Verfahren evaluieren soll.

Die Ergebnisse dieser Arbeit gehen in das GESIS-Projekt „Retrieval-Mehrwertdienste zur Weiterentwicklung wissenschaftlicher Fachportale wie vascoda und sowiport. Suchexpandierung und Re-Ranking“⁹⁴ (kurz IR-Mehrwertdienste) ein, das im Januar 2009 startet. Das Projekt wird in der Abteilung „Informationelle Prozesse in den Sozialwissenschaften“ durchgeführt (Leitung des Projekts: Jürgen Krause, Vivien Petras und Maximilian Stempfhuber).

Das Projekt IR-Mehrwertdienste fokussiert auf die Implementation und Evaluation von drei Mehrwertdiensten, die die Informationssuche in Digitalen Bibliotheken verbessern sollen.

Das Beispiel vascoda zeigt, dass Suchmaschinentechnologie (bei vascoda wird die Technologie von FAST eingesetzt) und das Konzept der semantischen Heterogenitätskomponenten noch nicht ausreichen, um den Überstieg in die heutige Welt polyzentrischer Informationsversorgung zu leisten. Zwei Schwachstellen, die auch bei Fachportalen und anderen Suchsystemen mit Crosskonkordanzen auftreten, lassen sich identifizieren.

a) Heterogenitätskomponenten schlüsseln automatisch zwischen verschiedenen Inhalterschließungsvokabularen um. Stringent funktioniert dies jedoch derzeit nur, wenn der Benutzer sich einer der kontrollierten Terminologien bedient, die durch Crosskonkordanzen miteinander vernetzt sind. In vielen Fällen wählt der Benutzer die Terme jedoch frei. Die Suchterme sind somit gar nicht oder nur zufällig einem der kontrollierten Vokabulare zugeordnet. Ein Mehrwertdienst, der frei gewählte Terme des Benutzers in Vorschläge aus dem kontrollierten Vokabular umformuliert, der STR (Search Term Recommender), kann die Treffergenauigkeit der Recherche für den Benutzer erhöhen (vgl. (Petras, 2006)).

b) Trotz der Heterogenitätskomponenten sind die Ergebnismengen in vielen Fällen noch zu hoch und die herkömmlichen Rankingverfahren – wie sie z. B. in FAST zur Verfügung stehen – greifen nur schlecht. Alternative Rankingverfahren (z. B. Bradfordizing oder Ranking nach Autorenzentralität) können daher die Rechervesituation bei den heutigen bibliographischen Datenbanken verbessern.

94 DFG-Projektkennziffer: INST 658/6-1; <http://www.gesis.org/irm>

Verallgemeinernd lassen sich folgende Vorteile der erwähnten Verfahren benennen:

- Als Instrument zum Vorschlagen besserer Suchterme reiht sich der STR in die Gruppe der Werkzeuge zur Suchverbesserung und Vermittlung zwischen Nutzer und Informationssystem ein. Durch eine Anfrageerweiterung mit Termen aus dem kontrollierten Wortschatz der inhaltlichen Beschreibungssysteme der Datenbanken (Thesauri, Schlagwortlisten, Klassifikationen) reduziert der STR die Heterogenität der eingesetzten Terminologien (durch Angleichung der Suchsprache an die kontrollierten Vokabulare) und verbessert damit die Vagheitsrelation V1 in Abbildung 3-3.
- Die Re-Rankingverfahren setzen auf einem beliebig zusammengesetzten Suchergebnis auf und nutzen die den Wissenschaften inhärenten Strukturen und Regelmäßigkeiten für die nachgeordnete Umsortierung (Re-Ranking) der Dokumente. Ziel dieser Dienste ist es, dem Nutzer eine alternative Sicht auf den Suchraum und das Suchergebnis zu liefern und damit den Suchprozess zu erweitern und zu verbessern.
- Da die Ansätze auf Kernzeitschriften und Kernautoren fokussieren, wird erwartet, dass diese Re-Rankingdienste auch zur Relevanzverbesserung der Ergebnisliste und somit zur Effizienzverbesserung der Suche insgesamt beitragen können.

Die Konzepte (STR & Re-Ranking) gewinnen eine besondere Qualität durch ihre Verbindung mit den Heterogenitätskomponenten, womit gleichzeitig Neuland beschritten wird. Sie setzen bewusst sowohl bei der Anfrage- als auch bei der Ergebnisseite an und können sich in einem iterativen Suchverfahren gegenseitig positiv beeinflussen (vgl. (Mayr et al., 2008)).

Abbildung 9-1 verdeutlicht den vorgeschlagenen Einsatz der Mehrwertdienste in den Phasen der Suche und das Zusammenspiel der einzelnen Komponenten.

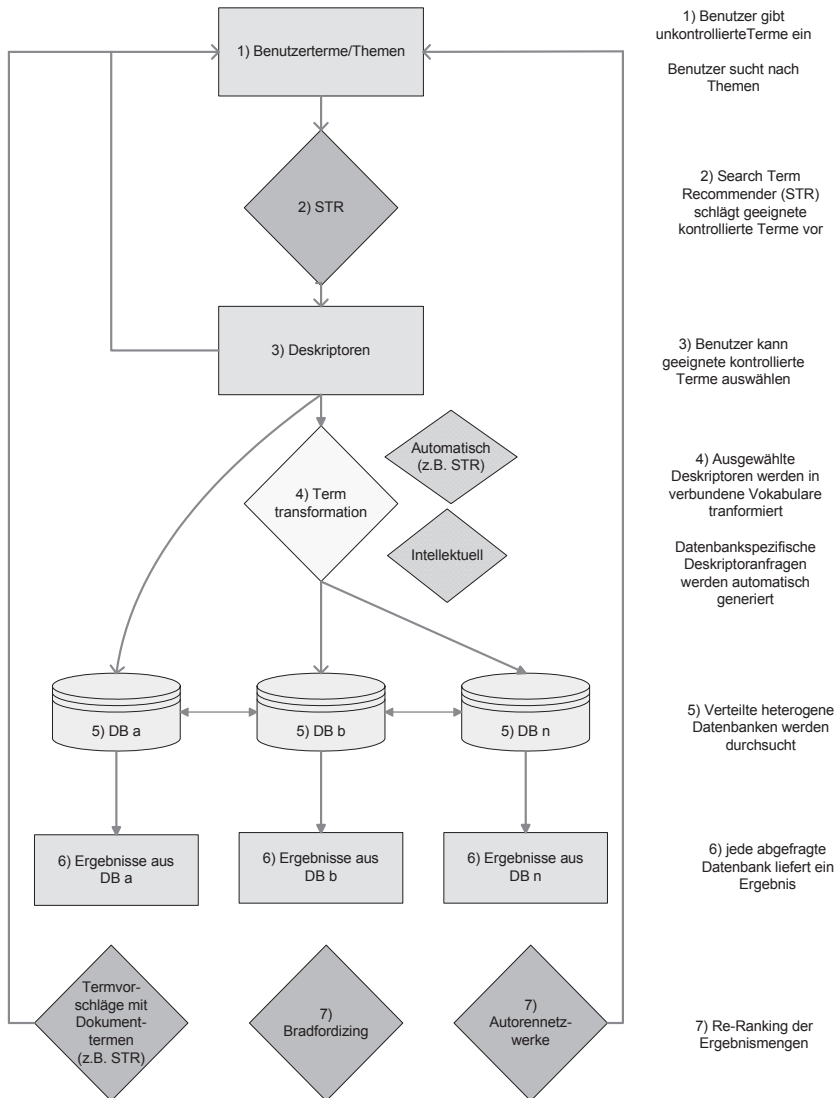


Abbildung 9-1: **Anwendung von Mehrwertdiensten zur Heterogenitätsbehandlung in den Phasen des Retrieval-Prozesses**

Um das Innovationspotential der drei Verfahren (STR, Bradfordizing und Autorennetzwerke) voll auszuschöpfen, bieten sich, über die isolierte Nutzung derselben hinaus, auch interessante Perspektiven für eine Kombination der Verfahren an, die im Papier ((Mayr et al., 2008)) konzeptuell angedacht wurden. Eine Koppelung der unterschiedlichen Verfahren/Mehrwertdienste führt mit großer Wahrscheinlichkeit zu weiteren Verbesserungen bei der Suche. Insbesondere die Behandlung von Benutzertermen (Vorschlag von kontrollierten Termen) gilt als vielversprechend.

“... the integration of three structural value adding methods which aim at reducing the semantic complexity represented in distributed DLs at several stages in the information retrieval process: query construction, search and ranking and re-ranking. The integration of the models will be done using Semantic Web technologies which should enhance further insights into the usage of these techniques. The intersection of the Semantic Web world with the DL world as mentioned in Krause (2008) ... will hopefully result in more sophisticated analytical tools and interfaces for the presentation of information adapted to users' needs.” ((Mayr et al., 2008) S. 221 - 222)

Ein weiterer Schwerpunkt des Projekts liegt in der Evaluation der einzelnen Verfahren. In diesem Zusammenhang sollten mindestens zwei Kategorien der Effektivität unterschieden werden. Zum einen besteht an jeden Retrieval-Dienst der Anspruch, die Suche insofern zu verbessern, so dass mehr relevante Dokumente für ein bestimmtes Informationsbedürfnis gefunden werden. Der Einfluss der neu aufgesetzten Mehrwertdienste in diesem Sinne lässt sich über traditionelle quantitative Retrieval-Evaluationsverfahren analysieren. Dazu sollen Nutzeranfragen und deren Suchresultate mit und ohne Hinzunahme der zusätzlichen Dienste verglichen und die Precision- und Recall-Werte gemessen werden.

Zusätzlich darf aber auch ein qualitativer Aspekt nicht außer Acht gelassen werden, der den Einfluss der Mehrwertdienste auf das Sucherlebnis an sich, d. h. die Zufriedenheit der Nutzer mit dem Ablauf der Suche untersucht. Dazu gehören Erfahrungen wie Schwierigkeiten in der Handhabung unterschiedlicher Retrieval-Instrumente und -oberflächen sowie die Gewissheit des Nutzers, mehrheitlich relevante und alle wichtigen relevanten Dokumente gefunden zu haben. Die erhöhte Schwierigkeit der Nutzerführung, Transparenz des Suchprozesses und der -ergebnisse sowie die Visualisierung der Terme bzw. neugeordneten Dokumente soll analysiert werden. Eine qualitative Analyse mittels Nutzerbefragungen kann darüber Aufschluss geben. Weiterhin müssen auch veränderte Einsatzszenarien für den integrierten Einsatz der Mehrwertdienste analysiert werden. Es muss geprüft werden, ob die Kombination der Mehrwertdienste weitere positive Effekte generieren kann oder, im Gegenteil, zusätzliche Vorkehrungen getroffen werden müssen, um potentielle negative Effekte auszuschalten.

10 Thesen

Nachfolgend werden die zentralen Thesen dieser Arbeit dargestellt. Diese Thesen sind als Ergebnisthesen formuliert, die die Ergebnisse der Experimente in dieser Arbeit zusammenfassen.

These 1: Das Konzept semantischer Heterogenitätsbehandlung (am Beispiel von Cross-konkordanzen) führt zu signifikanten Verbesserungen von Recall und Precision für das Information Retrieval mit kontrollierten Vokabularen. Der Recall für 7 getestete Datenbanken kann durchschnittlich um rund 92 % erhöht werden (das entspricht etwa einer Verdoppelung der relevanten Dokumente). Die Precision verbessert sich durchschnittlich um rund 53 %.

These 2: Bradford-Verteilungen treten erfahrungsgemäß in allen Fachgebieten und damit auch bei Recherchen in Datenbanken auf. Folglich kann das Bradfordizing zur Optimierung von Recherchen in der hier vorgeschlagenen Weise generell für themenspezifische Anfragen in Fachdatenbanken verwendet werden.

These 3: Die Anwendung informetrischer Verfahren zum Re-Ranking von Dokumenten vermag eine alternative Sicht auf Ergebnismengen zu ergeben, die vom Benutzer intuitiv als positiv empfunden wird. Positiv wird vor allem der Neuigkeits- und Erkenntnisgewinn bewertet, der durch hoch zitierte Arbeiten, Arbeiten zentraler Autoren, Artikel aus Kernzeitschriften und die Relevanzhierarchie des neuorganisierten Suchergebnisses entsteht.

These 4: Die Anwendung des Bradfordizings bzw. das Re-Ranking nach Kernzeitschriften für thematische Dokumentmengen führt zu signifikanten Verbesserungen der Precision zwischen den drei Zonen (Core, Zone 2 und Zone 3). Die Kernzeitschriften (Core) beinhalten signifikant mehr relevante Dokumente als Zeitschriften der Zone 2 oder der Zone 3. Der größte Precision-Gewinn ergibt sich zwischen Core und Zone 3-Zeitschriften.

These 5: Bradfordizing für thematisch konzentrierte Dokumentmengen lässt sich erfolgreich auf Monographien (Verlag als Selektionskriterium) übertragen. Die Anwendung des Bradfordizings bzw. das Re-Ranking nach Verlagen für Monographien führt zu geringeren Verbesserungen der Precision zwischen den drei Zonen (Core, Zone 2 und Zone 3). Der größte Precision-Gewinn ergibt sich auch hier wie bei den Zeitschriften zwischen Core und Zone 3-Verlagen.

These 6: Erstaunlicherweise zeigen die vorliegenden Ergebnisse, dass nach dem Bradfordizing Aufsätze in Kernzeitschriften häufiger als relevant für die Fragestellung bewertet werden, obwohl bei Bradford ja in allen Zonen nur relevante Dokumente Eingang finden. Erklärungsansätze für den Relevanzvorteil der Kernzeitschriften sind: a) die Tatsache, dass Kernzeitschriften in der Regel den State of the Art eines Themas repräsentieren, b) die Tatsache, dass Aufsätze in Kernzeitschriften, die das sog. peer-reviewing durchlaufen, von Spezialisten dieses Faches bewertet werden und nicht von eher Fachfremden und c) die Tatsache, dass bei den Publikationen im Kernbereich mehrere Aspekte des jeweiligen Forschungsgebietes abgedeckt werden.

Literaturverzeichnis

- Aiyepeku, Wilson O. (1977): The Bradford distribution theory: the compounding of Bradford periodical literatures in geography, *Journal of Documentation* 33 [3], Seite 210-219.
- Assem, Mark van; Malaisé, Véronique; Miles, Alistair und Schreiber, Guus (2006): A Method to Convert Thesauri to SKOS, *European Semantic Web Conference*. URL: <http://www.cs.vu.nl/~mark/papers/Assem06b.pdf>
- Bar-Ilan, Judit (1998/9): Search Engine Results over Time - A Case Study on Search Engine Stability, *Cybermetrics* [2/3]. URL: <http://www.cindoc.csic.es/cybermetrics/articles/v2i1p1.html>
- Bar-Ilan, Judit (2006): An ego-centric citation analysis of the works of Michael O. Rabin based on multiple citation indexes, *Information Processing & Management* 42 [6], Seite 1553-1566.
- Bar-Ilan, Judit (2008): Informetrics at the beginning of the 21st century - A review, *Journal of Informetrics* 2 [1], Seite 1-52.
- Bates, Marcia J. (1986): Subject Access in Online Catalogs: A Design Model, *Journal of the American Society for Information Science* 37 [6], Seite 357-376.
- Bates, Marcia J. (1989): The Design of Browsing and Berrypicking Techniques for the Online Search Interface, *Online Review* 13 [5], Seite 407-424. URL: <http://www.gseis.ucla.edu/faculty/bates/berrypicking.html>
- Bates, Marcia J. (1996): Document Familiarity, Relevance, and Bradford's Law: The Getty Online Searching Project Report No. 5, *Information Processing and Management* 32 [6], Seite 697-707.
- Bates, Marcia J. (1998): Indexing and Access for Digital Libraries and for the Internet: Human, Database, and Domain Factors, *Journal of the American Society for Information Science* 49 [13], Seite 1185-1205.
- Bates, Marcia J. (2002): Speculations on Browsing, Directed Searching, and Linking in Relation to the Bradford Distribution, Bruce, Harry; Fidel, Raya; Ingwersen, Peter und Vakkari, Pertti, *Fourth International Conference on Conceptions of Library and Information Science (CoLIS 4)* Seite 137-150. URL: http://www.gseis.ucla.edu/faculty/bates/articles/Searching_Bradford-m020430.html
- Beaver, D. deB. (2004): Does collaborative research have greater epistemic authority? *Scientometrics* 60 [3], Seite 399-408.
- Belew, Richard K. (2005): Scientific impact quantity and quality: Analysis of two sources of bibliographic data. URL: <http://arxiv.org/abs/cs.IR/0504036>
- Bennett, Rick; Hengel-Dittrich, Christina; O'Neill, Edward T. und Tillett, Barbara B. (2006): VIAF (Virtual International Authority File): Linking Die Deutsche Bibliothek and Library of Congress Name Authority Files, 72. IFLA Conference, Seoul, Korea. URL: <http://www.ifla.org/IV/ifla72/papers/123-Bennett-en.pdf>

- Bensman, S. J. (2001): Bradford's Law and fuzzy sets: statistical implications for library analyses, *IFLA Journal* 27 [4], Seite 238-246. URL: <http://www.garfield.library.upenn.edu/bensman/iflabradfordslawfuzzysets.pdf>
- Bergman, M. K. (2001): The Deep Web: Surfacing Hidden Value, *Journal of Electronic Publishing* 7 [1]. URL: <http://hdl.handle.net/2027/spo.3336451.0007.104>
- Binder, Gisbert; Marx, Jutta; Mutschke, Peter; Riege, Udo; Strötgen, Robert; Kokkelink, Stefan und Plümer, Judith (2002): Heterogenitätsbehandlung bei textueller Information verschiedener Datentypen und Inhaltserschließungsverfahren, *IZ-Arbeitsbericht; Nr. 24, IZ Sozialwissenschaften, Bonn*. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_24.pdf
- Binder, Gisbert; Stahl, Matthias und Faulborn, Lothar (2000): Vergleichsuntersuchung MESSENGER - FULCRUM: Projektbericht, *IZ-Arbeitsbericht; Nr. 18, IZ Sozialwissenschaften, Bonn*. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab18.pdf
- Binding, Ceri und Tudhope, Douglas (2004): KOS at your Service: Programmatic Access to Knowledge Organisation Systems, *Journal of Digital Information* 4 [4]. URL: <http://jodi.tamu.edu/Articles/v04/i04/Binding/>
- Björneborn, L. und Ingwersen, P. (2001): Perspective of webometrics, *Scientometrics* 50 [1], Seite 65-82.
- Blair, D. C. (1990): *Language and representation in information retrieval*, Elsevier Science Publishers, Amsterdam; New York.
- Blocks, Dorothee (2004): A qualitative study of thesaurus integration for end-user searching, Dissertation, University of Glamorgan/Prifysgol Morgannwg. URL: http://www.comp.glam.ac.uk/~FACET/dblocks/DBlocks_ThesisOnline_Main.html
- BMBF (2002): Information vernetzen – Wissen aktivieren, Seite 16. URL: http://www.dl-forum.de/dateien/Strategisches_Positionspapier.pdf
- Boekhorst, Peter te; Kayß, Matthias & Poll, Roswitha (2003): Nutzungsanalyse des Systems der überregionalen Literatur- und Informationsversorgung. URL: http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/ssg_bericht_teil_1.pdf
- Bonitz, Manfred (1980): Evidence for the invalidity of the Bradford Law for the single scientist, *Scientometrics* 2 [3], Seite 203-214.
- Bonitz, Manfred (1985): Journal ranking by selective impact. New method based on SDI results and journal impact factors, *Scientometrics* 7 [3-6], Seite 471-485.
- Bonitz, Manfred und Scharnhorst, Andrea (2001): Nicht alle Zeitschriften haben das gleiche Gewicht - Der harte Kern der Wissenschaftskommunikation, Discussion Paper FS II 01 - 307, Wissenschaftszentrum Berlin für Sozialforschung. URL: <http://skylla.wz-berlin.de/pdf/2001/ii01-307.pdf>
- Bookstein, A. (1990a): Informetric Distributions, Part I: Unified Overview, *JASIS* 41 [5], Seite 368-375. URL: http://topology.eecs.umich.edu/archive/bookstein_I.pdf

- Bookstein, A. (1990b): Informetric Distributions, Part II: Resilience to Ambiguity, JASIS 41 [5], Seite 376-386. URL: http://topology.eecs.umich.edu/archive/bookstein_II.pdf
- Bookstein, A. (1994): Towards a multi-disciplinary Bradford law, Scientometrics 30 [1], Seite 353-361.
- Borgman, Christine L. (1986): Why are online catalogs hard to use? Lessons learned from information-retrieval studies, JASIS 37 [6], Seite 387-400.
- Borgman, Christine L. (1996): Why are online catalogs still hard to use? JASIS 47 [7], Seite 493-503. URL: <http://www2.sims.berkeley.edu/courses/is202/f01/borgman.pdf>
- Bornmann, Lutz und Daniel, Hans-Dieter (2005): Does the h-index for ranking of scientists really work? Scientometrics 65 [3], Seite 391-392. URL: <http://www.lutz-bornmann.de/icons/h-index.pdf>
- Bortz, Jürgen (2005): Statistik für Human- und Sozialwissenschaftler, 6. Auflage, Springer Medizin Verlag, Heidelberg.
- Bradford, Samuel C. (1934): Sources of information on specific subjects, Engineering 137 [3550], Seite 85-86.
- Bradford, Samuel C. (1948): Documentation, Lockwood, London.
- Brin, Sergey und Page, Lawrence (1998): The Anatomy of a large-scale hypertextual Web search engine, Seventh International World Wide Web Conference, Brisbane, Australia. URL: <http://www-db.stanford.edu/~backrub/google.html>
- Brody, Timothy (2006): Evaluating Research Impact through Open Access to Scholarly Communication, Electronics and Computer Science, University of Southampton. URL: <http://eprints.ecs.soton.ac.uk/13313/>
- Brookes, B. C. (1968): The derivation and application of the Bradford-Zipf distribution, Journal of Documentation 24 [4], Seite 247-265.
- Brookes, B. C. (1969): Bradford's Law and the Bibliography of Science, Nature 224, Seite 953-956.
- Brookes, B. C. (1971): Optimum P% library of scientific periodicals, Nature 232 [5311], Seite 458-461.
- Brookes, B. C. (1977): Theory of the Bradford Law, Journal of Documentation 33 [3], Seite 180-209.
- Brophy, Jan und Bawden, David (2005): Is Google enough? Comparison of an internet search engine with academic library resources, Aslib Proceedings 57 [6], Seite 498-512.
- Buckland, Michael (1972): Are Obsolescence and Scattering related? Journal of Documentation 28 [3], Seite 242-246. URL: <http://www.sims.berkeley.edu/~buckland/obscat.html>
- Buckland, Michael (1999): Vocabulary as a Central Concept in Library and Information Science, Third International Conference on Conceptions of Library and Information Science. URL: <http://www.sims.berkeley.edu/~buckland/colisvoc.htm>

- Buckland, Michael und Gey, Fredric (1999): Mapping Entry Vocabulary to Unfamiliar Metadata Vocabularies, D-Lib Magazine 5 [1]. URL: <http://www.dlib.org/dlib/january99/buckland/01buckland.html>
- CARMEN-Projekt (2002): CARMEN - Abschlussbericht des Arbeitspakets 12 (AP 12) Crosskonkordanzen von Klassifikationen und Thesauri. URL: http://www.opus-bayern.de/uni-regensburg/volltexte/2003/242/pdf/CARMENAP12_Abschlussbericht_Netz.pdf
- Chan, Lois Mai und Zeng, Marcia Lei (2006): Metadata Interoperability and Standardization – A Study of Methodology Part I: Achieving Interoperability at the Schema Level, D-Lib Magazine 12 [6]. URL: <http://www.dlib.org/dlib/june06/chan/06chan.html>
- Chen, P.; Xie, H.; Maslov, S. und Redner, S. (2007): Finding scientific gems with Google's PageRank algorithm, Journal of Informetrics 1 [1], Seite 8-15.
- Chernov, Sergey; Fehling, Bernd; Kohlschütter, Christian; Nejd, Wolfgang; Pieper, Dirk und Summann, Friedrich (2006): Enabling Federated Search with Heterogeneous Search Engines: Combining FAST Data Search and Lucene, 73, 22. März 2006, <http://base.ub.uni-bielefeld.de/download/FedSearchReport.pdf>
- Chu, Heting (2003): Information Representation and Retrieval in the Digital Age, ASIST monograph series, Information Today, Medford, NJ, ISBN: 1-57387-172-9.
- Chu, Heting und Krichel, Thomas (2003): Current Awareness Service of the RePEc Digital Library, D-Lib Magazine 9 [12]. URL: <http://www.dlib.org/dlib/december03/chu/12chu.html>
- Chung, Y. K. (1994): Core International journals of Classification Systems: An application of Bradford's Law, Knowledge Organization 21 [2], Seite 75-83.
- Cimiano, Philipp; Pivk, Aleksander; Schmidt-Thieme, Lars und Staab, Steffen (2003): Learning Taxonomic Relations from Heterogeneous Sources of Evidence. URL: <http://www.uni-koblenz.de/~staab/Research/Publications/2005/OL-book-chapter-cimiano.pdf>
- Cline, Gloria S. (1981): Application of Bradford's law to citation data, College and Research Libraries 42 [1], Seite 53-61.
- Coleman, S. R. (1993): Bradford distributions of social-science bibliographies varying in definitional homogeneity, Scientometrics 27 [1], Seite 75-91.
- Coleman, S. R. (1994): Disciplinary variables that affect the shape of Bradford's bibliography, Scientometrics 29 [1], Seite 59-81.
- Dees, Werner (2008): Innovative Scientometric Methods for a Continuous Monitoring of Research Activities in Educational, Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting, Berlin. URL: <http://www.collnet.de/Berlin-2008/DeesWIS2008ism.pdf>
- Degkwitz, Andreas (2004): Bibliotheken im Kontext neuer Publikationsstrukturen, Bibliotheksdienst 38 [11], Seite 1417-1422. URL: http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte/heft9-1204/bibliotheken022204.pdf

- Dempsey, Lorcan (2006): The Library Catalogue in the New Discovery Environment: Some Thoughts, *Ariadne* [48]. URL: <http://www.ariadne.ac.uk/issue48/dempsey/>
- Depping, Ralf (2007): vascoda.de and the system of the German virtual subject libraries, Prasad, A. R. D. und Madalli, Devika P., *International Conference on Semantic Web & Digital Libraries (ICSD 2007)* Seite 304-314, Documentation Research & Training Centre, Indian Statistical Institute, Bangalore, India.
- Ding, Ying; Chowdhury, Gobinda G.; Foo, Schubert und Qian, Weizhong (2000): Bibliometric information retrieval system (BIRS): A web search interface utilizing bibliometric research results, *JASIS* 51 [13], Seite 1190-1204.
- Doerr, Martin (2001): Semantic Problems of Thesaurus Mapping, *Journal of Digital Information* 1 [8]. URL: <http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Doerr/>
- Doerr, Martin (2004): Semantic interoperability: Theoretical Considerations, Technical Report 345, ICS-FORTH.
- Efthimiadis, Efthimis N. (1996): Query Expansion, *Annual Review of Information Systems and Technology (ARIST)* 31, Seite 121-187. URL: <http://faculty.washington.edu/efthimis/pubs/Pubs/qe-arist/QE-arist.html>
- Egghe, Leo (2005): Expansion of the field of informetrics: Origins and consequences, *Information Processing and Management* 41 [6], Seite 1311-1316.
- Egghe, Leo und Rousseau, Ronald (1990): *Introduction to Informetrics: quantitative methods in library, documentation and information science*, Elsevier Science Publishers. URL: <http://eprints.rclis.org/archive/00003297/>
- Elsevier (2004): Scirus White Paper: How Scirus Works. URL: http://www.scirus.com/press/pdf/WhitePaper_Scirus.pdf
- Eto, H. (1988): Rising tail in Bradford Distribution: its interpretation and application, *Scientometrics* 13 [5-6], Seite 271-287.
- Eversberg, Bernhard (2002): *Zur Theorie der Bibliothekskataloge und Suchmaschinen*. URL: <http://www.allegro-c.de/formate/tks.htm>
- Ewert, Gisela und Umstätter, Walther (1997): *Lehrbuch der Bibliotheksverwaltung*, Hiersemann, Stuttgart, ISBN: 3-7772-9730-5.
- Eysenbach, Gunther (2006): Citation Advantage of Open Access Articles, *PLoS Biology* 4 [5]. URL: <http://dx.doi.org/10.1371/journal.pbio.0040157>
- Feijen, Martin; Horstmann, Wolfram; Manghi, Paolo; Robinson, Mary und Russell, Rosemary (2007): DRIVER: Building the Network for Accessing Digital Repositories across Europe, *Ariadne* 53. URL: <http://www.ariadne.ac.uk/issue53/feijen-et-al/>
- Fenske, Martin (2006): Modell eines automatisierbaren syntaktischen Metathesaurus und seine Eignung für parlamentarische Thesauri im Internet, Hauke, Petra und Umlauf, Konrad, *Vom Wandel der Wissensorganisation im Informationszeitalter – Festschrift für Walther Umstätter zum 65. Geburtstag*. URL: <http://edoc.hu-berlin.de/miscellanies/vom-27533/293/PDF/293.pdf>

- Fisher, Karen E.; Erdelez, Sandra und McKechnie, Lynne (2005): *Theory of Information Behavior*, ASIST Monograph Series, Information Today, Inc., Medford, New Jersey.
- Fournier, Johannes (2005): Zur Bedeutung von Open Access für das Publikationsverhalten DFG-geförderter Wissenschaftler, *ZfBB* [5], Seite 235-244. URL: http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/bericht_fournier.pdf
- Freyre, Elisabeth und Naudi, Max (2001): MACS: subject access across languages and networks, IFLA Satellite Meeting, Dublin, Ohio.
- Garfield, Eugene (1971): The Mystery of the Transposed Journal Lists - Wherein Bradford's Law of Scattering is Generalized according to Garfield's Law of Concentration, *Current Contents* 1 [17], Seite 222-223.
- Garfield, Eugene (1980): Bradford's Law and Related Statistical Patterns, *Current Contents* 4 [19], Seite 476-483. URL: <http://www.garfield.library.upenn.edu/essays/v4p476y1979-80.pdf>
- Garfield, Eugene (1990): How ISI Selects Journals for Coverage: Quantitative and Qualitative Considerations, *Current Contents* 13 [22], Seite 5-13. URL: <http://www.garfield.library.upenn.edu/essays/v13p185y1990.pdf>
- Garfield, Eugene (1996): The Significant Scientific Literature Appears In A Small Core Of Journals, *The Scientist* 10 [17], Seite 13. URL: [http://www.garfield.library.upenn.edu/commentaries/tsv10\(17\)p13y090296.html](http://www.garfield.library.upenn.edu/commentaries/tsv10(17)p13y090296.html)
- Geißelmann, Friedrich (1999): Zur dritten Auflage der RSWK, *Bibliotheksdienst* 33 [1], Seite 38-54. URL: http://bibliotheksdienst.zlb.de/1999/1999_01_Erschlie01.pdf
- Ginsparg, Paul (2003): Can Peer Review be better Focused? URL: <http://people.ccmr.cornell.edu/~ginsparg/blurb/pg02pr.html>
- Ginsparg, Paul (2007): Next-Generation Implications of Open Access, *CTWatch Quarterly* 3 [3]. URL: <http://www.ctwatch.org/quarterly/print.php?p=80>
- Godby, Carol Jean; Young, Jeffrey A. und Childress, Eric (2004): A Repository of Metadata Crosswalks, *D-Lib Magazine* 10 [12]. URL: <http://www.dlib.org/dlib/december04/godby/12godby.html>
- Goffman, W. und Warren, K. S. (1969): Dispersion of papers among journals based on a mathematical analysis of two diverse medical literatures, *Nature* 221 [5178], Seite 1205-1207.
- Hacker, Gerhard (2005): *Die Hybridbibliothek - Blackbox oder Ungeheuer?* Hacker, Gerhard und Seela, Torsten, *Bibliothek leben: Das deutsche Bibliothekswesen als Aufgabe für Wissenschaft und Politik / Festschrift für Engelbert Plassmann zum 70. Geburtstag*, Harrassowitz, Wiesbaden.
- Hage, Willem Robert van (2008): *Evaluating Ontology-Alignment Techniques*.
- Harman, Donna K. und Voorhees, Ellen M. (2006): TREC: An Overview, *Annual Review of Information Science and Technology* 40, Seite 113-156.

- Harnad, S. (1991): Post-Gutenberg Galaxy: The Fourth Revolution in the Means of Production of Knowledge, *Public-Access Computer Systems Review* 2 [1], Seite 39–53. URL: <http://cogprints.org/1580/>
- Harnad, S. und Brody, T. (2004): Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals, *D-Lib Magazine* 10 [6]. URL: <http://www.dlib.org/dlib/june04/harnad/06harnad.html>
- Harnad, Stevan; Brody, Tim; Vallières, François; Carr, Les; Hitchcock, Steve; Gingras, Yves; Oppenheim, Charles; Stamerjohanns, Heinrich und Hilf, Eberhard R. (2004): The green and the gold roads to Open Access, *Nature*. URL: <http://www.nature.com/nature/focus/accessdebate/21.html>
- Harter, Stephen P. (1996): Variations in Relevance Assessments and the Measurement of Retrieval Effectiveness, *JASIS* 47 [1], Seite 37–49. URL: http://www.asis.org/Publications/JASIS/Best_Jasist/1997Harter.pdf
- Havemann, Frank (2004): Eprints in der wissenschaftlichen Kommunikation. URL: <http://www.ib.hu-berlin.de/~fhavem/E-prints.pdf>
- Hawkins, Donald T (1978): The literature of noble gas compounds, *Journal of Chemical Information and Computer Sciences* 18 [4], Seite 190–199.
- Heine, M. H. (1998): Bradford ranking conventions and their application to a growing literature, *Journal of Documentation* 54 [3], Seite 303–331.
- Hellweg, Heiko; Krause, Jürgen; Mandl, Thomas; Marx, Jutta; Müller, Matthias N. O.; Mutschke, Peter und Strötgen, Robert (2001): Treatment of Semantic Heterogeneity in Information Retrieval, *IZ-Arbeitsbericht; Nr. 23, IZ Sozialwissenschaften, Bonn*. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_23.pdf
- Henzinger, Monika (2007): Search Technologies for the Internet, *Science* 317, Seite 468–471.
- Hill, L. L.; Buchel, O.; Janée, G. und Zeng, M. L. (2002): Integration of Knowledge Organization Systems into Digital Library Architectures. URL: <http://www.alexandria.ucsb.edu/~gjanee/archive/2002/kos-dl-paper.pdf>
- Hjørland, Birger und Nicolaisen, Jeppe (2005): Bradford's law of scattering: ambiguities in the concept of „subject“, 5th International Conference on Conceptions of Library and Information Science.
- Hobohm, Hans-Christoph (2004): Bibliotheken, Rainer Kuhlen, Thomas Seeger und Dietmar Strauch, *Grundlagen der praktischen Information und Dokumentation*, Seite 505–514, K. G. Saur, München.
- Hodge, Gail (2000): Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files, Washington, D.C. URL: <http://www.clir.org/pubs/reports/pub91/contents.html>

- Hood, William W. (1999): An informetric study of the distribution of bibliographic records in online databases: a case study using the literature of Fuzzy Set Theory (1965-1993), University of New South Wales, Sydney, Australia.
- Hood, William W. und Wilson, Concepción S. (2001): The literature of bibliometrics, scientometrics, and informetrics, *Scientometrics* 52 [2], Seite 291-314.
- Hood, William W. und Wilson, Concepción S. (2003): Informetric studies using databases: Opportunities and challenges, *Scientometrics* 58 [3], Seite 587-608.
- Hornbostel, Stefan (2006): Forschungsrankings: Artefakte oder Sichtbarkeit der Forschungsarbeit? Stempfhuber, Maximilian, In *die Zukunft publizieren. Herausforderungen an das Publizieren und die Informationsversorgung in den Wissenschaften*, Tagungsberichte, Band 11, Seite 263-277, Informationszentrum Sozialwissenschaften, Bonn.
- Ingwersen, Peter (1984): A cognitive view of three selected online search facilities, *Online Review* 8 [5], Seite 465-492.
- Ingwersen, Peter und Järvelin, Kalervo (2005): *The Turn - Integration of Information Seeking and Retrieval in Context*, Croft, W. Bruce, The Kluwer International Series on Information Retrieval, Springer, Dordrecht.
- Jacsó, Peter (2005): Google Scholar: the pros and the cons, *Online Information Review* 29 [2], Seite 208-214.
- Jacsó, Péter (2004): Citation searching, *Online Information Review* 28 [6], Seite 454-460.
- Jacsó, Péter (2008): Google Scholar revisited, *Online Information Review* 32 [1], Seite 102-114.
- Jin, Bihui und Wang, Bing (1999): Chinese science citation database: Its construction and application, *Scientometrics* 45 [2], Seite 325-332.
- Klas, Claus-Peter (2007): DAFFODIL: Strategische Unterstützung bei der Informationssuche in Digitalen Bibliotheken, Fachbereich Ingenieurwissenschaften, Universität Duisburg-Essen. URL: <http://www.is.informatik.uni-duisburg.de/bib/pdf/ir/Klas:07b.pdf>
- Kleinberg, Jon (1999): Authoritative Sources in a Hyperlinked Environment, *Journal of the ACM* 46 [5], Seite 604-632. URL: <http://portal.acm.org/citation.cfm?id=324140>
- Kling, R. und McKim, G. (1999): Scholarly Communication and the Continuum of Electronic Publishing, *JASIS* 50 [10], Seite 890-906. URL: <http://arxiv.org/abs/cs.CY/9903015>
- Kluck, Michael; Mandl, Thomas; Womser-Hacker, Christa (2002): Cross-Language Evaluation Forum (CLEF): Europäische Initiative zur Bewertung sprachübergreifender Retrievalverfahren, *Information - Wissenschaft und Praxis (nfd)* 53 [2], Seite 82-89.
- Koch, Traugott (2000): Quality-controlled subject gateways: definitions, typologies, empirical overview, *Online Information Review* 24 [1].
- Koch, Traugott; Neuroth, Heike und Day, Michael (2001): Renardus: Cross-browsing European subject gateways via a common classification system (DDC), IFLA Satellite Meeting, Dublin, Ohio. URL: <http://www.ukoln.ac.uk/metadata/renardus/papers/ifla-satellite/ifla-satellite.pdf>

- Kousha, Kayvan und Thelwall, Mike (2007): Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis, *Journal of the American Society for Information Science and Technology* 58 [7], Seite 1055-1065.
- Krause, Jürgen (1996): Informationserschließung und -bereitstellung zwischen Deregulation, Kommerzialisierung und weltweiter Vernetzung („Schalenmodell“), Bonn. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab6.pdf
- Krause, Jürgen (2003a): Standardisierung von der Heterogenität her denken: Zum Entwicklungsstand bilateraler Transferkomponenten für digitale Fachbibliotheken, IZ-Arbeitsbericht; Nr. 28, IZ Sozialwissenschaften, Bonn. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_28.pdf
- Krause, Jürgen (2003b): Standardization, heterogeneity and the quality of content analysis: a key conflict of digital libraries and its solution, IFLA 2003, World Library and Information Congress: 69th IFLA General Conference and Council, Berlin. URL: http://www.ifla.org/IV/ifla69/papers/085e_trans-Krause.pdf
- Krause, Jürgen (2004): Konkretes zur These, die Standardisierung von der Heterogenität her zu denken, *ZfBB: Zeitschrift für Bibliothekswesen und Bibliographie* 51 [2], Seite 76 - 89.
- Krause, Jürgen (2006): Shell Model, Semantic Web and Web Information Retrieval, Harms, Ilse; Luckhardt, Heinz-Dirk und Giessen, Hans W., *Information und Sprache: Beiträge zu Informationswissenschaft, Computerlinguistik, Bibliothekswesen und verwandten Fächern; Festschrift für Harald H. Zimmermann*, Seite 95 - 106, Saur, München.
- Krause, Jürgen (2008): Semantic heterogeneity: comparing new semantic web approaches with those of digital libraries, *Library Review* 57 [3], Seite 235-248.
- Krause, Jürgen (2008): Total Package Design für Digitale Bibliotheken und Fachinformation, Hutzler, Evelinde; Schröder, Albert und Schweikl, Gabriele, *Bibliotheken gestalten Zukunft. Kooperative Wege zur Digitalen Bibliothek. Dr. Friedrich Geißelmann zum 65. Geburtstag* Seite 185-206, Universitätsverlag Göttingen, Göttingen. URL: <http://epub.uni-regensburg.de/4564/>
- Krause, Jürgen; Mandl, Thomas und Stempfhuber, Maximilian (1997): Text-Fakten-Integration in ELVIRA, Bonn. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab12.pdf
- Krause, Jürgen und Mayr, Philipp (2006): Allgemeiner Bibliothekszugang und Varianten der Suchtypologie - Konsequenzen für die Modellbildung in vascoda, IZ-Arbeitsbericht Nr. 38, Informationszentrum Sozialwissenschaften, Bonn. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_38.pdf
- Krause, Jürgen; Niggemann, Elisabeth und Schwänzl, Roland (2003): Normierung und Standardisierung in sich verändernden Kontexten: Beispiel Virtuelle Fachbibliotheken, *ZfBB: Zeitschrift für Bibliothekswesen und Bibliographie* 50 [1], Seite 19 - 28.

- Krause, Jürgen und Stempfhuber, Maximilian (2005): Nutzerseitige Integration sozialwissenschaftlicher Text- und Dateninformationen aus verteilten Quellen, König, Christian; Stahl, Matthias und Wiegand, Erich, Datenfusion und Datenintegration: 6. Wissenschaftliche Jahrestagung, Seite 141 - 158, IZ Sozialwissenschaften, Bonn.
- Kurtz, Michael J.; Eichhorn, Guenther; Accomazzi, Alberto; Grant, Carolyn S.; Demleitner, Markus und Murray, Stephen S. (2005): The Effect of Use and Access on Citation, Information Processing and Management 41 [6], Seite 1395-1402. URL: <http://cfawww.harvard.edu/~kurtz/kurtz-effect.pdf>
- Kurtz, Michael J. und Henneken, Edwin A. (2007): Open Access does not increase citations for research articles from The Astrophysical Journal. URL: <http://front.math.ucdavis.edu/0709.0896>
- Lancaster, F. W. (1972): Vocabulary Control for Information Retrieval, Information Resources Press, Washington, D.C.
- Lancaster, F. W.; Gondek, Valerie; McCowan, Sherry und Reese, Cheri (1991): The relationship between literature scatter and journal accessibility in an academic special library, Collection Building 11 [1], Seite 19-22.
- Larson, Ray R.; McDonough, Jerome; O'Leary, Paul und Kuntz, Lucy (1996): Cheshire II: Designing a Next-Generation Online Catalog, JASIS 47 [7], Seite 555-567. URL: http://cheshire.sims.berkeley.edu/jasis_paper96.pdf
- Lauser, Boris; Johannsen, Gudrun; Caracciolo, Caterina; Keizer, Johannes; Hage, Willem Robert van und Mayr, Philipp (2008): Comparing human and automatic thesaurus mapping approaches in the agricultural domain, Greenberg, Jane und Klas, Wolfgang, Metadata for semantic and social applications: Proceedings of the 8. International Conference on Dublin Core and Metadata Applications, Seite 43-53, Univ.-Verl. Göttingen, Berlin. URL: <http://edoc.hu-berlin.de/docviews/abstract.php?lang=ger&id=29135>
- Lawrence, Steve (2001): Online or Invisible? Nature 411 [6837], Seite 521. URL: <http://citeseer.ist.psu.edu/online-nature01/>
- Lawrence, Steve; Giles, C. Lee und Bollacker, Kurt (1999): Digital Libraries and Autonomous Citation Indexing, IEEE Computer 32 [6], Seite 67-71. URL: <http://citeseer.ist.psu.edu/aci-computer/aci-computer99.html>
- Lawrence, Steve und Giles, Lee C. (1998): Searching the World Wide Web, Science [280]. URL: <http://clgiles.ist.psu.edu/papers/Science-98.pdf>
- Lawrence, Steve und Giles, Lee C. (1999): Accessibility of information on the web, Nature 400, Seite 107-109. URL: <http://nike.psu.edu/classes/ist597/2003-fall/papers/web-access.pdf>
- Lewandowski, Dirk (2005): Web Information Retrieval: Technologien zur Informationssuche im Internet 7, DGI, Frankfurt am Main. URL: <http://www.durchdenken.de/lewandowski/web-ir/>
- Lewandowski, Dirk (2008): Handbuch Internet-Suchmaschinen, Akademische Verlagsgesellschaft Aka.

- Lewandowski, Dirk und Mayr, Philipp (2006): Exploring the academic invisible web, *Library Hi Tech* 24 [4], Seite 529-539. URL: <http://www.ib.hu-berlin.de/~mayr/arbeiten/LHT-2006.pdf>
- Liang, A. C. und Sini, M. (2006): Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, tools, procedures, *New Review in Hypermedia and Multimedia* 12 [1], Seite 51-62.
- Lin, Jimmy (2008): PageRank without hyperlinks: Reranking with PubMed related article networks for biomedical text retrieval, *BMC Bioinformatics* 9. URL: <http://www.biomedcentral.com/1471-2105/9/270>
- Lockett, Mary W. (1989): The Bradford distribution. A review of the literature, 1934-1987, *Library and Information Science Research* 11 [1], Seite 21-36.
- Lossau, Norbert (2004): Search Engine Technology and Digital Libraries - Libraries Need to Discover the Academic Internet, *D-Lib Magazine* 10 [6]. URL: <http://www.dlib.org/dlib/june04/lossau/06lossau.html>
- Lotka, Alfred (1926): The frequency distribution of scientific productivity, *Journal of the Washington Academy of Sciences* 16 [12], Seite 317-323.
- Macgregor, George; Joseph, Anu und Nicholson, Dennis (2007): A SKOS Core approach to implementing an M2M terminology mapping server, *International Conference on Semantic Web and Digital Libraries (ISCD-2007)*, Bangalore, India. URL: <http://strathprints.strath.ac.uk/2970/1/strathprints002970.pdf>
- Marchionini, Gary (1995): *Information seeking in electronic environments*, Cambridge series on human-computer interaction; 9, Cambridge University Press, Cambridge; New York.
- Marx, Matthias N. O. (2005): *Empirische Ergebnisse zur Evaluation semantischer Transformationen*, unveröffentlichter IZ-Arbeitsbericht, IZ Sozialwissenschaften, Bonn.
- Mayr, Philipp (2006): Informationsangebote für das Wissenschaftsportal vascoda - eine Bestandsaufnahme, IZ-Arbeitsbericht Nr. 37, Informationszentrum Sozialwissenschaften, Bonn. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_37.pdf
- Mayr, Philipp (2006): Thesauri, Klassifikationen & Co – die Renaissance der kontrollierten Vokabulare? Hauke, Petra und Umlauf, Konrad, *Vom Wandel der Wissensorganisation im Informationszeitalter. Festschrift für Walther Umstätter zum 65. Geburtstag*, Seite 151-170, Bock + Herchen Verlag, Bad Honnef. URL: <http://edoc.hu-berlin.de/miscellanies/vom-27533/151/PDF/151.pdf>
- Mayr, Philipp (2007): Integrität und Integration von elektronischen Publikationen - Modellüberlegungen im Umfeld der Digitalen Bibliothek, Havemann, Frank; Parthey, Heinrich und Umstätter, Walther, *Wissenschaftsforschung Jahrbuch 2007*, Seite 107-119, Gesellschaft für Wissenschaftsforschung, Berlin. URL: <http://edoc.hu-berlin.de/miscellanies/wifo2007/PDF/wifo2007-107-119.pdf>
- Mayr, Philipp (2008): An evaluation of Bradfordizing effects, *Proceedings of WIS 2008*, Berlin, Fourth International Conference on Webometrics, Informetrics and Sciento-

- metrics & Ninth COLLNET Meeting, Humboldt-Universität zu Berlin. URL: <http://www.collnet.de/Berlin-2008/MayrWIS2008ebe.pdf>
- Mayr, Philipp; Mutschke, Peter und Petras, Vivien (2008): Reducing semantic complexity in distributed digital libraries: Treatment of term vagueness and document re-ranking, *Library Review* 57 [3], Seite 213-224. URL: <http://www.emeraldinsight.com/10.1108/00242530810865484>
- Mayr, Philipp und Petras, Vivien (2008a): Building a terminology network for search: the KoMoHe project, Greenberg, Jane und Klas, Wolfgang, *Metadata for semantic and social applications: Proceedings of the 8. International Conference on Dublin Core and Metadata Applications*, Seite 177-182, Univ.-Verl. Göttingen, Berlin. URL: <http://edoc.hu-berlin.de/docviews/abstract.php?lang=ger&id=29148>
- Mayr, Philipp und Petras, Vivien (2008b): Cross-concordances: terminology mapping and its effectiveness for information retrieval, *IFLA World Library and Information Congress*, Québec, Canada. URL: http://www.ifla.org/IV/ifla74/papers/129-Mayr_Petras-en.pdf
- Mayr, Philipp und Tosques, Fabio (2005a): Webometrische Analysen mit Hilfe der Google Web APIs, *Information - Wissenschaft & Praxis* 56 [1], Seite 41-48. URL: http://www.ib.hu-berlin.de/~mayr/arbeiten/Mayr_Tosques_IWP05.pdf
- Mayr, Philipp und Tosques, Fabio (2005b): Google Web APIs - An Instrument for Webometric Analyses? Ingwersen, Peter und Larsen, Birger, *Proceedings of the 10th International Conference of the International Society for Scientometrics and Informetrics*, Seite 677-678, Karolinska University Press, Stockholm, Sweden. URL: http://www.ib.hu-berlin.de/~mayr/arbeiten/ISSI2005_Mayr_Toques.pdf
- Mayr, Philipp und Umstätter, Walther (2007): Why is a new Journal of Informetrics needed? *Cybermetrics* 11 [1]. URL: <http://www.cindoc.csic.es/cybermetrics/articles/v11i1p1.html>
- Mayr, Philipp und Umstätter, Walther (2008): Eine bibliometrische Zeitschriftenanalyse zu JoI, Scientometrics und NfD bzw. IWP, *Information - Wissenschaft & Praxis* 59 [6-7], Seite 353-360.
- Mayr, Philipp und Walter, Anne-Kathrin (2007): An exploratory study of Google Scholar, *Online Information Review* 31 [6], Seite 814-830. URL: <http://www.ib.hu-berlin.de/~mayr/arbeiten/OIR31-6.pdf>
- Mayr, Philipp und Walter, Anne-Kathrin (2008): Mapping Knowledge Organization Systems, Ohly, H. Peter; Netscher, Sebastian und Mitgutsch, Konstantin, *Fortschritte der Wissensorganisation*, Band 10. Kompatibilität, Medien und Ethik in der Wissensorganisation, Seite 80-95, Ergon, Würzburg. URL: <http://www.ib.hu-berlin.de/~mayr/arbeiten/Mayr-Walter-ISK006.pdf>
- McCown, Frank und Nelson, Michael L. (2007): Search Engines and Their Public Interfaces: Which APIs are the Most Synchronized? *16th International World Wide Web Conference*, Banff, Canada. URL: <http://www2007.org/posters/poster868.pdf>

- McVeigh, E. Marie (2004): Open Access Journals in the ISI Citation Databases: Analysis of Impact Factors and Citation Patterns. A citation study from Thomson Scientific. URL: <http://www.thomsonscientific.com/media/presentrep/essayspdf/openaccesscitations2.pdf>
- Miles, Alistair (2006): SKOS: Requirements for Standardization, International Conference on Dublin Core and Metadata Applications. URL: <http://isegserv.itd.rl.ac.uk/public/skos/press/dc2006/camera-ready-paper.pdf>
- Mizzaro, Stefano (1997): Relevance: The whole history, *Journal of the American Society for Information Science* 48 [9], Seite 810-832.
- Moed, H. F. (2006): The effect of ‚Open Access‘ upon citation impact: An analysis of ArXiv’s Condensed Matter Section. URL: <http://arxiv.org/abs/cs.DL/0611060>
- Müller, Matthias N. O. (2004): Fachinformationslandschaft in Deutschland, IZ-Arbeitsmaterial Nr.16 (unveröffentlicht), Informationszentrum Sozialwissenschaften, Bonn.
- Mutschke, Peter (2003): Mining Networks and Central Entities in Digital Libraries: A Graph Theoretic Approach applied to Co-Author Networks; Posterpräsentation, IDA 2003 – The 5th International Symposium on Intelligent Data Analysis, Berlin. URL: <http://fuzzy.cs.uni-magdeburg.de/confs/ida2003/>
- Mutschke, Peter (2004a): Autorennetzwerke: Netzwerkanalyse als Mehrwertdienst für Informationssysteme, *Information zwischen Kultur und Marktwirtschaft*. Proceedings des 9. Internationalen Symposiums für Informationswissenschaft (ISI 2004), Chur. URL: <http://www.inf-wiss.uni-konstanz.de/infwiss/download/cc-isi04-art9.pdf>
- Mutschke, Peter (2004b): Autorennetzwerke: Verfahren der Netzwerkanalyse als Mehrwertdienste für Informationssysteme, IZ-Arbeitsbericht Nr. 32, Informationszentrum Sozialwissenschaften, Bonn. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_32.pdf
- Nacke, Otto (1979): Informetrie: ein neuer Name für eine neue Disziplin, *Nachrichten für Dokumentation* 30 [6], Seite 219-226.
- Nentwich, Michael (2003): *Cyberscience: research in the age of the Internet*, Austrian Academy of Sciences Press, Vienna, ISBN: 3-7001-3188-7.
- Newman, M. E. J. (2005): Power laws, Pareto distributions and Zipf’s law, *Contemporary Physics* 46 [5], Seite 323-351. URL: <http://arxiv.org/abs/cond-mat/0412004>
- Nicolaisen, Jeppe und Hjørland, Birger (2007): Practical potentials of Bradford’s law: A critical examination of the received view, *Journal of Documentation* 63 [3], Seite 359-377.
- Nikolai, Ralf (2002): *Thesaurusföderationen: Ein Rahmenwerk für die flexible Integration von heterogenen, autonomen Thesauri*, Dissertation, Universität Fridericiana zu Karlsruhe, Hilden (Rheinland). URL: www.ubka.uni-karlsruhe.de/indexer-vvv/2003/informatik/6

- Nourmohammadi, H. A. und Umstätter, W. (2004): Die Verteilung der Autorenhäufigkeit in wissenschaftlichen Zeitschriften bei verschiedenen Themen und Ländern, *Information Wissenschaft und Praxis* 55 [5], Seite 275-281.
- Nunzio, Giorgio Maria Di und Ferro, Nicola (2005): DIRECT: A System for Evaluating Information Access Components of Digital Libraries, *ECDL* 2005.
- Oluic-Vukovic, V. (1997): Bradford's distribution: from the classical bibliometric 'law' to the more general stochastic models, *Journal of the American Society for Information Science* 48 [9], Seite 833-842.
- Page, Larry; Brin, Sergey; Motwani, R. und Winograd, T. (1998): The PageRank Citation Ranking: Bringing Order to the Web. URL: <http://citeseer.ist.psu.edu/page98pagerank.html>
- Panzer, Michael (2008): Semantische Integration heterogener und unterschiedlichsprachiger Wissensorganisationssysteme: CrissCross und jenseits, Ohly, H. Peter; Netscher, Sebastian und Mitgutsch, Konstantin, *Fortschritte in der Wissensorganisation*, Band 10. Kompatibilität, Medien und Ethik in der Wissensorganisation, Seite 61-69, *Ergon*, Würzburg.
- Patel, Manjula; Koch, Traugott; Doerr, Martin und Tsinaraki, Chrisa (2005): Semantic Interoperability in Digital Library Systems. URL: <http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DLs/>
- Peritz, Bluma C. (1990): A Bradford distribution for bibliometrics, *Scientometrics* 18 [5-6], Seite 323-329.
- Petras, Vivien (2006): *Translating Dialects in Search: Mapping between Specialized Languages of Discourse and Documentary Languages*, University of California, Berkeley, Berkeley, USA. URL: <http://www.sims.berkeley.edu/~vivienp/diss/>
- Petras, Vivien; Bärtsch, Stefan und Stempfhuber, Maximilian (2007): The Domain-Specific Track at CLEF 2007, *CLEF 2007 Workshop*, Budapest, Hungary. URL: http://www.clef-campaign.org/2007/working_notes/petras_overview-ds-CLEF2007.pdf
- Pianos, Tamara (2005): Was macht vascoda? Vision und Wirklichkeit, *Zeitschrift für Bibliothekswesen und Bibliographie* 52 [2], Seite 67-78.
- Plaunt, C. und Norgard, B. A. (1998): An association-based method for automatic indexing with a controlled vocabulary, *Journal of the American Society for Information Science* 49 [10], Seite 888-902.
- Pontigo, J. und Lancaster, F. W. (1986): Qualitative aspects of the Bradford distribution, *Scientometrics* 9 [1-2], Seite 59-70.
- Price, Derek de Solla (1976): A general theory of bibliometric and other cumulative advantage processes, *Journal of the American Society for Information Science* 27, Seite 292-306.
- Qiu, Liwen und Tague, Jean (1990): Complete or incomplete data sets. The Groos droop investigated, *Scientometrics* 19 [3-4], Seite 223-237.

- Rao, Ravichandra I. K. (1998): An analysis of Bradford multipliers and a model to explain law of scattering, *Scientometrics* 41 [1-2], Seite 93-100.
- Rousseau, Ronald (1997): Sitations: an exploratory study, *Cybermetrics* 1. URL: <http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html>
- Saracevic, Tefko (1975): Relevance: A review of and a framework for the thinking on the notion in information science, *Journal of the American Society for Information Science* 26 [6], Seite 321-343.
- Saracevic, Tefko (2007a): Relevance: A review of and a framework for the thinking on the notion in information science. Part II: Nature and Manifestations of Relevance, *Journal of the American Society for Information Science and Technology* 58 [13], Seite 1915-1933.
- Saracevic, Tefko (2007b): Relevance: A review of and a framework for the thinking on the notion in information science. Part III: Behavior and Effects of Relevance, *Journal of the American Society for Information Science and Technology* 58 [13], Seite 2126-2144.
- Schwarz, Iris und Umstätter, Walther (1999): Die vernachlässigten Aspekte des Thesaurus: dokumentarische, pragmatische, semantische und syntaktische Einblicke, *Nachrichten für Dokumentation* 50 [4], Seite 197-203. URL: <http://www.ib.hu-berlin.de/~wumsta/infopub/pub1996f/pub110.html>
- Seeger, Thomas (2004): Ein Blick voraus in die Vergangenheit: Von Visionen und Konzepten der frühen Dokumentationsbewegung, Hammwöhner, Rainer; Rittberger, Marc und Semar, Wolfgang, *Wissen in Aktion. Der Primat der Pragmatik als Motto der Konstanzer Informationswissenschaft. Festschrift für Rainer Kuhlen*, Seite 181-193, UVK Verlagsgesellschaft mbH, Konstanz. URL: http://www.inf-wiss.uni-konstanz.de/infwiss/download/festschrift/cc-festschrift_RK-art14.pdf
- Seiffert, Florian (2001): Eine Analyse der Verbunddaten des HBZ, *ABI-Technik* 21 [2], Seite 125-146.
- Sen, S. K. und Chatterjee, S. K. (1998): Bibliographic scattering and time: an empirical study through temporal partitioning of bibliographies, *Scientometrics* 41 [1-2], Seite 134-154.
- Sperber, Dan und Wilson, Deirdre (1995): *Relevance: Communication and Cognition*, Second Edition. Blackwell, Oxford.
- Stein, Lincoln (2002): Creating a bioinformatics nation, *Nature* 417 [6888], Seite 119-120. URL: <http://www.nature.com/nature/journal/v417/n6885/full/417119a.html>
- Stempfhuber, Max (2003): *Objektorientierte Dynamische Benutzungsoberflächen ODIN: Behandlung semantischer und struktureller Heterogenität in Informationssystemen mit den Mitteln der Softwareergonomie*, Forschungsberichte; 6, IZ Sozialwissenschaften, Bonn, ISBN: 3-8206-0142-2.
- Stempfhuber, Max (2008): Die Rolle von „open access“ im Rahmen des wissenschaftlichen Publizierens, Alexander von Humboldt-Stiftung, Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen (Diskussionspapiere der Alexander

- von Humboldt-Stiftung Nr. 12) Seite 66-81, Bonn. URL: <http://www.humboldt-foundation.de/pls/web/docs/F1316/publikationsverhalten.pdf>
- Strötgen, Robert (2007): Anfragetransfers zur Integration von Internetquellen in Digitalen Bibliotheken auf der Grundlage statistischer Termrelationen. URL: <http://www.opus-bayern.de/bib-info/volltexte/2007/291/pdf/2007-bid-leipzig-stroetgen-asemos-paper.pdf>
- Summann, Friedrich und Lossau, Norbert (2004): Search Engine Technology and Digital Libraries - Moving from Theory to Practice, D-Lib Magazine 10 [9]. URL: <http://www.dlib.org/dlib/september04/lossau/09lossau.html>
- Svenonius, Elaine (2000): The intellectual foundation of information organization, Digital libraries and electronic publishing, MIT Press, Cambridge, Mass., ISBN: 0262194333 (hc alk. paper).
- Swan, Alma und Brown, Sheridan (2005): Open access self-archiving: An author study. URL: <http://cogprints.org/4385/>
- Tapaswi, M. P. und Maheswarappa, B. S. (1999): Ranking serials in oceanography: an analysis based on the indian contributions and their citations, Scientometrics 44 [1], Seite 93-127.
- Taylor, Arlene G. (1999): The organization of information, Library and information science text series, Libraries Unlimited, Englewood, Colo.
- Tenopir, Carol (1982): Distribution of citations in databases in a multidisciplinary field, Online Review 6 [5], Seite 399-419.
- Thelwall, M.; Vaughan, L. und Björneborn, L. (2005): Webometrics, Annual Review of Information Science and Technology (ARIST) 39, Seite 81-135. URL: http://vip.db.dk/lb/papers/Thelwall_et_al_2005_Webometrics_ARIST.pdf
- Thelwall, Mike (2007): Extracting Accurate and Complete Results from Search Engines: Case Study Windows Live, JASIST 59 [1], Seite 38-50.
- Tosques, Fabio und Mayr, Philipp (2005): Web Services - Einsatzmöglichkeiten für das Information Retrieval im WWW, 27. DGI-Online-Tagung, Frankfurt am Main. URL: http://www.ib.hu-berlin.de/~mayr/arbeiten/tosques_mayr_dgi05.pdf
- Tosques, Fabio und Mayr, Philipp (2008): Programmierschnittstellen der kommerziellen Suchmaschinen, Lewandowski, Dirk, Handbuch Internet-Suchmaschinen, Seite 116-147, Akademische Verlagsanstalt AKA, Heidelberg. URL: http://www.ib.hu-berlin.de/~mayr/arbeiten/tosques-mayr_Handbuch08.pdf
- Triolo, V. A. und Bao, D. (1993): A decision model for technical journal deselection with an experiment in biomedical communications, Journal of the American Society for Information Science 44 [3], Seite 148-160.
- Tsay, Ming-Yueh; Jou, Shioh-Jen und Ma, Sheau-Shin (2000): A Bibliometric Study of Semiconductor Literature, 1978-1997, Scientometrics 49 [3], Seite 491-509.

- Tudhope, D.; Koch, T. und Heery, R. (2006): Terminology Services and Technology: JISC state of the art review, Environment, JISC Information. URL: http://www.jisc.ac.uk/Terminology_Services_and_Technology_Review_Sep_06
- Tudhope, Douglas; Binding, Ceri; Blocks, Dorothee und Cunliffe, Daniel (2006): Query expansion via conceptual distance in thesaurus indexed collections, *Journal of Documentation* 62 [4], Seite 509-533.
- Umstätter, Walter (2005): Bibliographie, Kataloge, Suchmaschinen - Das Ende der Dokumentation als modernes Bibliothekswesen, *Bibliotheksdienst* 39 [11], Seite 1442-1456. URL: http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte2005/Bibliotheken011105.pdf
- Umstätter, Walther (1991): Wäre es nicht langsam Zeit, die Informationstechnologie in der bibliothekarischen Sacherschließung etwas ernster zu nehmen? Ein Wort zur RSWK, *ABI-Technik* 11 [4]. URL: <http://www.ib.hu-berlin.de/~wumsta/infopub/pub1991f/pub64.html>
- Umstätter, Walther (1999): Bibliothekswissenschaft als Teil der Wissenschaftswissenschaft - unter dem Aspekt der Interdisziplinarität, Umstätter, W. und Wessel, K.-F., *Interdisziplinarität - Herausforderung an die Wissenschaftlerinnen und Wissenschaftler*, Seite 146-160, Kleine Verlag, Bielefeld. URL: <http://www.ib.hu-berlin.de/~wumsta/lectg.html>
- Umstätter, Walther (2003): Was ist und was kann eine wissenschaftliche Zeitschrift heute und morgen leisten, Parthey, Heinrich und Umstätter, Walther, *Wissenschaftliche Zeitschrift und Digitale Bibliothek: Wissenschaftsforschung Jahrbuch 2002*, Seite 143-166. URL: http://www.ib.hu-berlin.de/~wumsta/infopub/pub2001f/JB02_143-166.pdf
- Umstätter, Walther (2004): Szientometrische Verfahren, Rainer Kuhlen, Thomas Seeger und Dietmar Strauch, *Grundlagen der praktischen Information und Dokumentation* Seite 237-243, K. G. Saur, München.
- Umstätter, Walther (2005): Anmerkungen zu Birger Hjørland und Jeppe Nicolaisen: Bradford's Law of Scattering: Ambiguities in the Concept of „Subject“, *LIBREAS* [3]. URL: http://www.ib.hu-berlin.de/~libreas/libreas_neu/ausgabe3/008ums.htm
- Umstätter, Walther und Rehm, M. (1984): *Bibliothek und Evolution*, *Nachrichten für Dokumentation* 35 [6], Seite 237-249. URL: <http://www.ib.hu-berlin.de/~wumsta/pub27.html>
- Umstätter, Walther und Wagner-Döbler, Roland (2005): *Einführung in die Katalogkunde: Vom Zettelkatalog zur Suchmaschine*, 3., vollständig neu überarbeitete Auflage, Hiersemann, Stuttgart.
- van Rijsbergen, C.V. (1979): *Information Retrieval*, 2nd Edition, Butterworth, London. URL: <http://www.dcs.gla.ac.uk/Keith/Preface.html>
- vascoda (2006): *Core-Set für eine geisteswissenschaftlich und/oder regional orientierte Virtuelle Fachbibliothek*.
- Vickery, Brian C. (1948): Bradford's law of scattering, *Journal of Documentation* 4 [3], Seite 198-203.

- Vizine-Goetz, Diane; Hickey, Carol; Houghton, Andrew und Thompson, Roger (2004): Vocabulary Mapping for Terminology Services, *Journal of Digital Information* 4 [4]. URL: <http://jodi.tamu.edu/Articles/v04/i04/Vizine-Goetz/>
- von Ungern-Sternberg, S. (2000): Bradford's law in the context of information provision, *Scientometrics* 49 [1], Seite 161-186.
- Voorhees, Ellen M. (2007): TREC: Continuing information retrieval's tradition of experimentation, *Communications of the ACM* 50 [11], Seite 51-54.
- Voß, Jakob (2007): Tagging, Folksonomy & Co - Renaissance of Manual Indexing? 10th International Symposium for Information Science (ISI 2007), Köln. URL: <http://arxiv.org/abs/cs/0701072>
- Wagner-Döbler, Roland (1996): Two components of a causal explanation of Bradford's Law, *Journal of Information Science* 22 [2], Seite 125-132.
- Wagner-Döbler, Roland (1997): Time dependencies of Bradford distributions: structures of journal output in Twentieth century logic and Nineteenth century mathematics, *Scientometrics* 39 [3], Seite 231-252.
- Wasserman, Stanley und Faust, Katherine (1994): *Social Network Analysis. Methods and Applications, Structural Analysis in the Social Sciences*, 8, Cambridge University Press.
- White, Howard D. (1981): ‚Bradfordizing‘ search output: how it would help online users, *Online Review* 5 [1], Seite 47-54.
- White, Howard D. (2005): On Extending Informetrics: An Opinion Paper, *Proceedings of the 10th International Conference of the International Society for Scientometrics and Informetrics*, Seite 442-449, Karolinska University Press, Stockholm, Sweden.
- White, Howard D. (2007a): Combining bibliometrics, information retrieval, and relevance theory, Part 1: First examples of a synthesis, *JASIST* 58 [4], Seite 536-559.
- White, Howard D. (2007b): Combining bibliometrics, information retrieval, and relevance theory, Part 2: Some implications for information science, *JASIST* 58 [4], Seite 583-605.
- Wilkinson, E. A. (1972): The ambiguity of Bradford's law, *Journal of Documentation* 28 [2], Seite 122-130.
- Williams, Martha E. (2006): *The state of databases today: 2006*, Gale Directory of Databases, Gale Group, Detroit.
- Wilson, Concepción S. (1995): The formation of subject literature collections for bibliometric analysis: the case of the topic of Bradford's Law of Scattering, University of New South Wales, Sydney, Australia. URL: <http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN1999.0056>
- Wilson, Concepción S. (1998): Defining subject collections for informetric analyses: The effect of varying the subject aboutness level, *Scientometrics* 41 [1-2], Seite 209-223.
- Wilson, Concepción S. (1999): Informetrics, *Annual Review of Information Science and Technology (ARIST)* 34, Seite 107-247.

- Wolff, Christian (2007): Überlegungen zur Weiterentwicklung von vascoda, vascoda Strategiepapier.
- Wolfram, Dietmar (2003): Applied informetrics for information retrieval research, Libraries Unlimited, Westport, CT.
- Womser-Hacker, Christa (2004): Theorie des Information Retrieval III: Evaluierung, Rainer Kuhlen, Thomas Seeger und Dietmar Strauch (Hrsg.), Grundlagen der praktischen Information und Dokumentation, Seite 227-235, K. G. Saur, München.
- Worthen, D. B. (1975): The application of Bradford's law to monographs, Journal of Documentation 31 [1], Seite 19-25.
- Yablonsky, A. I. (1980): On fundamental regularities of the distribution of scientific productivity, Scientometrics 2 [1], Seite 3-34.
- Zeng, Marcia Lei und Chan, Lois Mai (2004): Trends and Issues in Establishing Interoperability Among Knowledge Organization Systems, Journal of the American Society for Information Science and Technology 55 [3], Seite 377-395.
- Zhang, Xueying (2006): Concept integration of document databases using different indexing languages, Information Processing and Management 42 [1], Seite 121-135.

Anhang

Beispielfragebogen

Kurzinterview¹ zu Kernzeitschriften am Beispiel von Fragestellungen der Evaluationsinitiative CLEF

Topic: Geschlecht und Karrierechancen

Nr.: 193

Datum:

[Begrüßung, Vorbemerkung und kurze Einführung zum Ziel des Interviews, Erklärung: Was ist eine Kernzeitschrift]

Die folgenden Fragen beziehen sich auf das von Ihnen vorgeschlagene Topic **[Geschlecht und Karrierechancen]**. Ich lese Ihnen nun die Topic-Beschreibung vor.

Titel des Topics: Geschlecht und Karrierechancen

Beschreibung des Topics: Suchen Sie Dokumente zur geschlechtsspezifischen Karriereentwicklung.

Ausführlichere Beschreibung: Relevante Dokumente beinhalten Informationen zur Frage unterschiedlicher Karriereentwicklungen von Männern und Frauen, insbesondere mit Bezug auf die verschiedenen Berufssparten und Karrierestufen. Von Interesse sind auch Diskussionen über individuelle Merkmale, die als karriereförderlich oder karrierehinderlich identifiziert werden.

Dieses Topic wurde im Rahmen der CLEF-Evaluation 2007 berücksichtigt. Die gefundenen Dokumente wurden intellektuell nach Relevanz begutachtet.

Eine weitere Analyse dieses Topics hat die Zeitschriften identifiziert, die (im Zeitraum 1990 - 2000) die meisten Artikel zu dem Thema veröffentlicht haben. Diese Zeitschriften sollen nachfolgend Kernzeitschriften genannt werden.

¹ Erklärungen finden sich in eckigen Klammern und sind kursiv gedruckt.

Frage 1: Nennen Sie mir bitte spontan mehrere Zeitschriften, die Sie zu den Kernzeitschriften (häufigsten Zeitschriften) dieses Themas in der Datenbank SOLIS zählen.

[Zeitschriften mitprotokollieren oder KEINE vermerken]

Ich lese Ihnen nun die analysierten Kernzeitschriften dieses Themas vor.

[Vorlesen der Liste der Kernzeitschriften]

- Zeitschrift fuer Frauenforschung
- Frauenforschung Informationsdienst des Forschungsinstituts Frau und Gesellschaft
- Koelner Zeitschrift fuer Soziologie und Sozialpsychologie
- Work, Employment and Society a journal of the British Sociological Association
- American journal of sociology
- Aus Politik und Zeitgeschichte Beilage zur Wochenzeitung Das Parlament
- Zeitschrift fuer Sozialisationsforschung und Erziehungssoziologie
- Feministische Studien
- Das Argument Zeitschrift fuer Philosophie und Sozialwissenschaften
- WSI Mitteilungen Monatszeitschrift des Wirtschafts- und Sozialwissenschaftlichen Instituts in der Hans-Boeckler-Stiftung
- Journal of occupational behaviour

Frage 2: Hätten Sie diese Zeitschriften als häufigste Veröffentlichungsorgane für Ihr Topic erwartet?

Antworten:

1) ja

2) z.T. wenn 2) Welche hätten Sie erwartet? Welche hätten Sie nicht erwartet?

3) nein wenn 3) Welche hätten Sie nicht erwartet?

4) weiß nicht

[Antworten notieren]

Frage 3: Können Sie sich vorstellen, dass die eben vorgelesenen Kernzeitschriften relevantere Arbeiten zum Thema publizieren als die übrigen in SOLIS ausgewerteten Zeitschriften?

Antworten:

1) ja wenn 1) warum?

2) nein wenn 2) warum?

3) weiß nicht

[Antworten notieren]

Frage 4: Welche Zeitschriften, die regelmäßig Literatur zu den obigen Topics beinhalten, vermissen Sie in dieser Liste? Zählen Sie bitte mehrere Zeitschriftentitel auf!

[Zeitschriften mitprotokollieren oder KEINE vermerken]

Signifikanztests

Signifikanztest CLEF 2003–2007 (Zeitschriftenartikel)

Tabelle 57: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 125 Topics der Evaluation CLEF 2003–2007. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|-----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 125 | ,24667 | ,02206 |
| | Pz3 | 125 | ,16714 | ,01495 |
| Paaren 2 | Pcore | 125 | ,24667 | ,02206 |
| | Pz2 | 125 | ,21593 | ,01931 |
| Paaren 3 | Pz2 | 125 | ,21593 | ,01931 |
| | Pz3 | 125 | ,16714 | ,01495 |
| Paaren 4 | Pcore | 125 | ,24667 | ,02206 |
| | Pbaseline | 125 | ,19566 | ,01750 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 125 | ,802 | ,000 |
| Paaren 2 | Pcore & Pz2 | 125 | ,830 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 125 | ,828 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 125 | ,944 | ,000 |

Tabelle 58: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Phaseline) für die 125 Topics der Evaluation CLEF 2003–2007. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | | | | | | | | |
|--------------------------------|-------------------|--------|--------------------|--------|---------------------------------|--------|--------------------------------------|-------|-----------------|
| Gepaarte Differenzen | | | | | | | | | |
| | Mittelwert | | Standardabweichung | | Standardfehler des Mittelwertes | | 95% Konfidenzintervall der Differenz | | Sig. (2-seitig) |
| | Untere | Obere | Untere | Obere | Untere | Obere | Untere | Obere | |
| Paaren 1 | Pcore - Pz3 | ,09416 | ,15043 | ,01345 | ,12079 | ,06753 | 6,998 | 124 | ,000 |
| Paaren 2 | Pcore - Pz2 | ,04584 | ,13796 | ,01234 | ,07026 | ,02142 | 3,715 | 124 | ,000 |
| Paaren 3 | Pz2 - Pz3 | ,04832 | ,12172 | ,01089 | ,06987 | ,02677 | 4,438 | 124 | ,000 |
| Paaren 4 | Pcore - Phaseline | ,04816 | ,08936 | ,00799 | ,06398 | ,03234 | 6,026 | 124 | ,000 |

| Statistik für Test(b) | | | | |
|--------------------------------------|-------------|-------------|-----------|-------------------|
| | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 | Phaseline - Pcore |
| Z | -6,309(a) | -3,622(a) | -4,177(a) | -5,443(a) |
| Asymptotische Signifikanz (2-seitig) | ,000 | ,000 | ,000 | ,000 |

Signifikanztest CLEF 2005–2007 (Monographien)

Tabelle 59: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Phaseline) für die 75 Topics der Evaluation CLEF 2005–2007. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Monographien.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 75 | ,22181 | ,02561 |
| | Pz3 | 75 | ,18455 | ,02131 |
| Paaren 2 | Pcore | 75 | ,22181 | ,02561 |
| | Pz2 | 75 | ,20074 | ,02318 |
| Paaren 3 | Pz2 | 75 | ,20074 | ,02318 |
| | Pz3 | 75 | ,18455 | ,02131 |
| Paaren 4 | Pcore | 75 | ,22181 | ,02561 |
| | Pbaseline | 75 | ,18145 | ,02095 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 75 | ,751 | ,000 |
| Paaren 2 | Pcore & Pz2 | 75 | ,809 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 75 | ,815 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 75 | ,861 | ,000 |

Tabelle 60: T-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Phaseline) für die 75 Topics der Evaluation CLEF 2005–2007. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Monographien.

| Test bei gepaarten Stichproben | | | | | | | | | | | | | | |
|--------------------------------|-------------------|----------------------|-------------------------|------------------------------------|---|--------|------------|-------------------------|------------------------------------|--------|-------|----|-----------------|--|
| | | Gepaarte Differenzen | | | | | | | | T | | df | Sig. (2-seitig) | |
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | Untere | Obere | | | |
| | | Untere | Obere | Untere | Obere | Untere | Obere | | | | | | | |
| Paaren 1 | Pcore - Pz3 | ,02347 | ,14751 | ,01703 | -,01047 | ,05740 | 1,378 | 74 | | | ,172 | | | |
| Paaren 2 | Pcore - Pz2 | ,02573 | ,13196 | ,01524 | -,00463 | ,05610 | 1,689 | 74 | | | ,095 | | | |
| Paaren 3 | Pz2 - Pz3 | -,00227 | ,11818 | ,01365 | -,02946 | ,02492 | -,166 | 74 | | | ,869 | | | |
| Paaren 4 | Pcore - Phaseline | ,02053 | ,11318 | ,01307 | -,00551 | ,04657 | 1,571 | 74 | | | ,120 | | | |

Signifikanztest CLEF 2003 (Zeitschriftenartikel)

Tabelle 61: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2003. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,25047 | ,05009 |
| | Pz3 | 25 | ,14721 | ,02944 |
| Paaren 2 | Pcore | 25 | ,25047 | ,05009 |
| | Pz2 | 25 | ,22106 | ,04421 |
| Paaren 3 | Pz2 | 25 | ,22106 | ,04421 |
| | Pz3 | 25 | ,14721 | ,02944 |
| Paaren 4 | Pcore | 25 | ,25047 | ,05009 |
| | Palljour | 25 | ,19660 | ,03932 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 25 | ,869 | ,000 |
| Paaren 2 | Pcore & Pz2 | 25 | ,914 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 25 | ,814 | ,000 |
| Paaren 4 | Pcore & Palljour | 25 | ,977 | ,000 |

Tabelle 62: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2003. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | Gepaarte Differenzen | | | | | | T | df | Sig. (2-seitig) |
|--------------------------------|------------------|---|------------------------------------|--------|--------|--------|-------|----|------|-----------------|
| | | 95% Konfidenzintervall der Differenz | | | | | | | | |
| | | Standard- abweichung | Standardfehler des Mittelwertes | Obere | Untere | | | | | |
| Paaren 1 | Pcore - Pz3 | ,13680 | ,14262 | ,02852 | ,07793 | ,19567 | 4,796 | 24 | ,000 | |
| Paaren 2 | Pcore - Pz2 | ,07600 | ,10202 | ,02040 | ,03389 | ,11811 | 3,725 | 24 | ,001 | |
| Paaren 3 | Pz2 - Pz3 | ,06080 | ,13251 | ,02650 | ,00610 | ,11550 | 2,294 | 24 | ,031 | |
| Paaren 4 | Pcore - Palljour | ,07200 | ,07165 | ,01433 | ,04243 | ,10157 | 5,025 | 24 | ,000 | |

| Statistik für Test(b) | | | |
|---|-------------|-------------|-----------|
| | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 |
| Z | -3,974(a) | -3,257(a) | -2,016(a) |
| Asymptotische Signifikanz (2-seitig) | ,000 | ,001 | ,044 |
| | | | ,000 |

Signifikanztest CLEF 2004 (Zeitschriftenartikel)

Tabelle 63: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2004. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,25128 | ,05026 |
| | Pz3 | 25 | ,14992 | ,02998 |
| Paaren 2 | Pcore | 25 | ,25128 | ,05026 |
| | Pz2 | 25 | ,19746 | ,03949 |
| Paaren 3 | Pz2 | 25 | ,19746 | ,03949 |
| | Pz3 | 25 | ,14992 | ,02998 |
| Paaren 4 | Pcore | 25 | ,25128 | ,05026 |
| | Pbaseline | 25 | ,18449 | ,03690 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 25 | ,836 | ,000 |
| Paaren 2 | Pcore & Pz2 | 25 | ,783 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 25 | ,760 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 25 | ,951 | ,000 |

Tabelle 64: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2004. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | Gepaarte Differenzen | | | | | | |
|---|-------------------|----------------------|-------------------------|------------------------------------|---|-------------------|-------|------|
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | T | df |
| | | | | | Obere | Untere | | |
| Paaren 1 | Pcore - Pz3 | ,09280 | ,15051 | ,03010 | ,03067 | ,15493 | 3,083 | 24 |
| Paaren 2 | Pcore - Pz2 | ,04120 | ,15621 | ,03124 | -,02328 | ,10568 | 1,319 | 24 |
| Paaren 3 | Pz2 - Pz3 | ,05160 | ,12834 | ,02567 | -,00138 | ,10458 | 2,010 | 24 |
| Paaren 4 | Pcore - Pbaseline | ,04640 | ,09478 | ,01896 | ,00728 | ,08552 | 2,448 | 24 |
| Statistik für Test(b) | | | | | | | | |
| Z | Pz3 - Pcore | -2,793(a) | Pz2 - Pcore | -1,113(a) | Pz3 - Pz2 | Pbaseline - Pcore | | |
| | | | | | -1,559(a) | -1,989(a) | | |
| Asymptotische Signifikanz (2-seitig) | | ,005 | | ,266 | ,119 | | | ,047 |

Signifikanztest CLEF 2005 (Zeitschriftenartikel)

Tabelle 65: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2005. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,29484 | ,05897 |
| | Pz3 | 25 | ,16737 | ,03347 |
| Paaren 2 | Pcore | 25 | ,29484 | ,05897 |
| | Pz2 | 25 | ,24010 | ,04802 |
| Paaren 3 | Pz2 | 25 | ,24010 | ,04802 |
| | Pz3 | 25 | ,16737 | ,03347 |
| Paaren 4 | Pcore | 25 | ,29484 | ,05897 |
| | Pbaseline | 25 | ,22310 | ,04462 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 25 | ,870 | ,000 |
| Paaren 2 | Pcore & Pz2 | 25 | ,908 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 25 | ,886 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 25 | ,971 | ,000 |

Tabelle 66: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2005. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | Gepaarte Differenzen | | | | | | Sig. (2-seitig) | |
|--------------------------------|-------------------|----------------------|-------------------------|------------------------------------|---|--------|-------|-----------------|------|
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | T | | |
| | | | | | Obere | Untere | | | |
| Paaren 1 | Pcore - Pz3 | ,13600 | ,17059 | ,03412 | ,06559 | ,20641 | 3,986 | 24 | ,001 |
| Paaren 2 | Pcore - Pz2 | ,06920 | ,12639 | ,02528 | ,01703 | ,12137 | 2,738 | 24 | ,011 |
| Paaren 3 | Pz2 - Pz3 | ,06680 | ,12023 | ,02405 | ,01717 | ,11643 | 2,778 | 24 | ,010 |
| Paaren 4 | Pcore - Pbaseline | ,07000 | ,09491 | ,01898 | ,03082 | ,10918 | 3,688 | 24 | ,001 |

| Statistik für Test(b) | | | | | |
|-----------------------|---|-------------|-------------|-----------|-------------------|
| Z | Asymptotische Signifikanz (2-seitig) | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 | Pbaseline - Pcore |
| | | -3,176(a) | -2,330(a) | -2,604(a) | -2,967(a) |
| | | ,001 | ,020 | ,009 | ,003 |

Signifikanztest CLEF 2006 (Zeitschriftenartikel)

Tabelle 67: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2006. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,23604 | ,04721 |
| | Pz3 | 25 | ,19214 | ,03843 |
| Paaren 2 | Pcore | 25 | ,23604 | ,04721 |
| | Pz2 | 25 | ,22272 | ,04454 |
| Paaren 3 | Pz2 | 25 | ,22272 | ,04454 |
| | Pz3 | 25 | ,19214 | ,03843 |
| Paaren 4 | Pcore | 25 | ,23604 | ,04721 |
| | Pbaseline | 25 | ,20329 | ,04066 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 25 | ,860 | ,000 |
| Paaren 2 | Pcore & Pz2 | 25 | ,768 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 25 | ,872 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 25 | ,932 | ,000 |

Tabelle 68: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2006. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | Gepaarte Differenzen | | | | | | |
|--------------------------------|-------------------|----------------------|-------------------------|------------------------------------|---|--------|-------|----|
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | T | df |
| | | | | | Obere | Untere | | |
| Paaren 1 | Pcore - Pz3 | ,04280 | ,12088 | ,02418 | -,00710 | ,09270 | 1,770 | 24 |
| Paaren 2 | Pcore - Pz2 | ,02040 | ,15683 | ,03137 | -,04434 | ,08514 | ,650 | 24 |
| Paaren 3 | Pz2 - Pz3 | ,02240 | ,10914 | ,02183 | -,02265 | ,06745 | 1,026 | 24 |
| Paaren 4 | Pcore - Pbaseline | ,02200 | ,08727 | ,01745 | -,01402 | ,05802 | 1,260 | 24 |

| Statistik für Test(b) | | | | |
|---|--|-------------|-------------|-------------------|
| | | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 |
| | | -1,827(a) | -,958(a) | -1,142(a) |
| | | | | Pbaseline - Pcore |
| | | | | -1,341(a) |
| Asymptotische Signifikanz (2-seitig) | | ,068 | ,338 | ,253 |
| | | | | ,180 |

Signifikanztest CLEF 2007 (Zeitschriftenartikel)

Tabelle 69: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2007. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,20436 | ,04087 |
| | Pz3 | 25 | ,16398 | ,03280 |
| Paaren 2 | Pcore | 25 | ,20436 | ,04087 |
| | Pz2 | 25 | ,20349 | ,04070 |
| Paaren 3 | Pz2 | 25 | ,20349 | ,04070 |
| | Pz3 | 25 | ,16398 | ,03280 |
| Paaren 4 | Pcore | 25 | ,20436 | ,04087 |
| | Pbaseline | 25 | ,17218 | ,03444 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 25 | ,686 | ,000 |
| Paaren 2 | Pcore & Pz2 | 25 | ,757 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 25 | ,801 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 25 | ,896 | ,000 |

Tabelle 70: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2007. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | | | | | | | | |
|---|-------------------|---|-------------------------|------------------------------------|-------------------|--------|-------|----|-----------------|
| Gepaarte Differenzen | | | | | | | | | |
| | | 95% Konfidenzintervall der Differenz | | | | | | | |
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | Obere | Untere | T | df | Sig. (2-seitig) |
| Paaren 1 | Pcore - Pz3 | ,06240 | ,15062 | ,03012 | ,00023 | ,12457 | 2,071 | 24 | ,049 |
| Paaren 2 | Pcore - Pz2 | ,02240 | ,14205 | ,02841 | -,03623 | ,08103 | ,788 | 24 | ,438 |
| Paaren 3 | Pz2 - Pz3 | ,04000 | ,12186 | ,02437 | -,01030 | ,09030 | 1,641 | 24 | ,114 |
| Paaren 4 | Pcore - Pbaseline | ,03040 | ,09154 | ,01831 | -,00738 | ,06818 | 1,661 | 24 | ,110 |
| Statistik für Test(b) | | | | | | | | | |
| | | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 | Pbaseline - Pcore | | | | |
| Z | | -2,242(a) | -,772(a) | -1,994(a) | -1,874(a) | | | | |
| Asymptotische Signifikanz (2-seitig) | | ,025 | ,440 | ,046 | ,061 | | | | |

Signifikanztest CLEF 2005 (Monographien)

Tabelle 71: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2005. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Monographien.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,20396 | ,04079 |
| | Pz3 | 25 | ,19276 | ,03855 |
| Paaren 2 | Pcore | 25 | ,20396 | ,04079 |
| | Pz2 | 25 | ,18243 | ,03649 |
| Paaren 3 | Pz2 | 25 | ,18243 | ,03649 |
| | Pz3 | 25 | ,19276 | ,03855 |
| Paaren 4 | Pcore | 25 | ,20396 | ,04079 |
| | Pbaseline | 25 | ,18616 | ,03723 |

| Korrelationen bei gepaarten Stichproben | | | |
|---|-------------------|-------------|-------------|
| | N | Korrelation | Signifikanz |
| Paaren 1 | Pcore & Pz3 | ,865 | ,000 |
| Paaren 2 | Pcore & Pz2 | ,833 | ,000 |
| Paaren 3 | Pz2 & Pz3 | ,892 | ,000 |
| Paaren 4 | Pcore & Pbaseline | ,923 | ,000 |

Tabelle 72: T-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Phaseline) für die 25 Topics der Evaluation CLEF 2005. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf **Monographien**.

| Test bei gepaarten Stichproben | | | | | | | | | |
|--------------------------------|-------------------|------------|-------------------------|------------------------------------|---|--------|--------|----|-----------------|
| Gepaarte Differenzen | | | | | | | | | |
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | T | df | Sig. (2-seitig) |
| | | | | | Obere | Untere | | | |
| Paaren 1 | Pcore - Pz3 | ,01720 | ,10358 | ,02072 | -,02556 | ,05996 | ,830 | 24 | ,415 |
| Paaren 2 | Pcore - Pz2 | ,04280 | ,11367 | ,02273 | -,00412 | ,08972 | 1,883 | 24 | ,072 |
| Paaren 3 | Pz2 - Pz3 | -,02560 | ,08771 | ,01754 | -,06180 | ,01060 | -1,459 | 24 | ,157 |
| Paaren 4 | Pcore - Phaseline | ,01600 | ,07832 | ,01566 | -,01633 | ,04833 | 1,022 | 24 | ,317 |

Signifikanztest CLEF 2006 (Monographien)

Tabelle 73: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2006. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Monographien.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,22716 | ,04543 |
| | Pz3 | 25 | ,17640 | ,03528 |
| Paaren 2 | Pcore | 25 | ,22716 | ,04543 |
| | Pz2 | 25 | ,21400 | ,04280 |
| Paaren 3 | Pz2 | 25 | ,21400 | ,04280 |
| | Pz3 | 25 | ,17640 | ,03528 |
| Paaren 4 | Pcore | 25 | ,22716 | ,04543 |
| | Pbaseline | 25 | ,17227 | ,03445 |

| Korrelationen bei gepaarten Stichproben | | | |
|---|-------------------|-------------|-------------|
| | N | Korrelation | Signifikanz |
| Paaren 1 | Pcore & Pz3 | ,635 | ,001 |
| Paaren 2 | Pcore & Pz2 | ,902 | ,000 |
| Paaren 3 | Pz2 & Pz3 | ,766 | ,000 |
| Paaren 4 | Pcore & Pbaseline | ,822 | ,000 |

Tabelle 74: T-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Phaseline) für die 25 Topics der Evaluation CLEF 2006. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf **Monographien**.

| Test bei gepaarten Stichproben | | Gepaarte Differenzen | | | | | |
|--------------------------------|-------------------|----------------------|--------------------|---------------------------------|--------------------------------------|--------|-----------------|
| | | Mittelwert | Standardabweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | Sig. (2-seitig) |
| | | | | | Obere | Untere | |
| Paaren 1 | Pcore - Pz3 | ,04800 | ,17844 | ,03569 | -,02566 | ,12166 | 1,345 |
| Paaren 2 | Pcore - Pz2 | ,00200 | ,09870 | ,01974 | -,03874 | ,04274 | ,101 |
| Paaren 3 | Pz2 - Pz3 | ,04600 | ,13814 | ,02763 | -,01102 | ,10302 | 1,665 |
| Paaren 4 | Pcore - Phaseline | ,03320 | ,13015 | ,02603 | -,02052 | ,08692 | 1,275 |
| | | | | | | | 24 |
| | | | | | | | 24 |
| | | | | | | | 24 |
| | | | | | | | 24 |

Signifikanztest CLEF 2007 (Monographien)

Tabelle 75: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2007. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Monographien.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 25 | ,23501 | ,04700 |
| | Pz3 | 25 | ,18774 | ,03755 |
| Paaren 2 | Pcore | 25 | ,23501 | ,04700 |
| | Pz2 | 25 | ,19374 | ,03875 |
| Paaren 3 | Pz2 | 25 | ,19374 | ,03875 |
| | Pz3 | 25 | ,18774 | ,03755 |
| Paaren 4 | Pcore | 25 | ,23501 | ,04700 |
| | Pbaseline | 25 | ,18759 | ,03752 |

| Korrelationen bei gepaarten Stichproben | | | |
|---|----|-------------|-------------|
| | N | Korrelation | Signifikanz |
| Paaren 1 | 25 | ,759 | ,000 |
| Paaren 2 | 25 | ,688 | ,000 |
| Paaren 3 | 25 | ,827 | ,000 |
| Paaren 4 | 25 | ,842 | ,000 |

Tabelle 76: T-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 25 Topics der Evaluation CLEF 2007. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf **Monographien**.

| Test bei gepaarten Stichproben | | Gepaarte Differenzen | | | | | |
|--------------------------------|-------------------|----------------------|-------------------------|------------------------------------|---|--------|-----------------|
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | Sig. (2-seitig) |
| | | | | | Obere | Untere | |
| Paaren 1 | Pcore - Pz3 | ,00520 | ,15330 | ,03066 | -,05808 | ,06848 | ,170 |
| Paaren 2 | Pcore - Pz2 | ,03240 | ,17348 | ,03470 | -,03921 | ,10401 | ,934 |
| Paaren 3 | Pz2 - Pz3 | -,02720 | ,11245 | ,02249 | -,07362 | ,01922 | -1,209 |
| Paaren 4 | Pcore - Pbaseline | ,01240 | ,12716 | ,02543 | -,04009 | ,06489 | ,488 |

Signifikanztest KoMoHe (Zeitschriftenartikel)

Tabelle 77: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 39 Topics der Evaluation KoMoHe. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 39 | ,27492 | ,04402 |
| | Pz3 | 39 | ,22475 | ,03599 |
| Paaren 2 | Pcore | 39 | ,27492 | ,04402 |
| | Pz2 | 39 | ,24875 | ,03983 |
| Paaren 3 | Pz2 | 39 | ,24875 | ,03983 |
| | Pz3 | 39 | ,22475 | ,03599 |
| Paaren 4 | Pcore | 39 | ,27492 | ,04402 |
| | Pbaseline | 39 | ,24163 | ,03869 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 39 | ,901 | ,000 |
| Paaren 2 | Pcore & Pz2 | 39 | ,927 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 39 | ,950 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 39 | ,966 | ,000 |

Tabelle 78: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 39 Topics der Evaluation KoMoHe. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf Zeitschriftenartikel. (a basiert auf positiven Rängen. b Wilcoxon-Test)

| Test bei gepaarten Stichproben | | | | | | | | | |
|---|-------------------|-------------|-------------------------|------------------------------------|---|--------|-------|----|-----------------|
| Gepaarte Differenzen | | | | | | | | | |
| | | Mittelwert | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | T | df | Sig. (2-seitig) |
| | | | | | Obere | Untere | | | |
| Paaren 1 | Pcore - Pz3 | ,06026 | ,12167 | ,01948 | ,02082 | ,09970 | 3,093 | 38 | ,004 |
| Paaren 2 | Pcore - Pz2 | ,03231 | ,10345 | ,01657 | -,00123 | ,06584 | 1,950 | 38 | ,059 |
| Paaren 3 | Pz2 - Pz3 | ,02795 | ,07828 | ,01253 | ,00257 | ,05332 | 2,230 | 38 | ,032 |
| Paaren 4 | Pcore - Pbaseline | ,03128 | ,07537 | ,01207 | ,00685 | ,05571 | 2,592 | 38 | ,013 |
| Statistik für Test(b) | | | | | | | | | |
| | | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 | Pbaseline - Pcore | | | | |
| Z | | -2,682(a) | -1,982(a) | -2,096(a) | -2,508(a) | | | | |
| Asymptotische Signifikanz (2-seitig) | | ,007 | ,047 | ,036 | ,012 | | | | |

Signifikanztest KoMoHe (Monographien)

Tabelle 79: Statistische Werte und Korrelationen für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 30 Topics der Evaluation KoMoHe. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf Monographien.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | Pcore | 30 | ,26127 | ,04770 |
| | Pz3 | 30 | ,23606 | ,04310 |
| Paaren 2 | Pcore | 30 | ,26127 | ,04770 |
| | Pz2 | 30 | ,22849 | ,04172 |
| Paaren 3 | Pz2 | 30 | ,22849 | ,04172 |
| | Pz3 | 30 | ,23606 | ,04310 |
| Paaren 4 | Pcore | 30 | ,26127 | ,04770 |
| | Pbaseline | 30 | ,23800 | ,04345 |

| Korrelationen bei gepaarten Stichproben | | | | |
|---|-------------------|-------------|-------------|------|
| | N | Korrelation | Signifikanz | |
| Paaren 1 | Pcore & Pz3 | 30 | ,939 | ,000 |
| Paaren 2 | Pcore & Pz2 | 30 | ,922 | ,000 |
| Paaren 3 | Pz2 & Pz3 | 30 | ,940 | ,000 |
| Paaren 4 | Pcore & Pbaseline | 30 | ,968 | ,000 |

Tabelle 80: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (Pcore, Pz2, Pz3 und Pbaseline) für die 30 Topics der Evaluation KoMoHe. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf **Monographien**. (a basiert auf positiven Rängen. b Basiert auf negativen Rängen. c Wilcoxon-Test)

| Test bei gepaarten Stichproben | | | | | | | | | |
|--------------------------------|-------------------|-------------------------|------------------------------------|---|---------|--------|-------|-----------------|------|
| Gepaarte Differenzen | | | | | | | | | |
| | | Standard- abweichung | Standardfehler des Mittelwertes | 95% Konfidenzintervall der Differenz | | T | df | Sig. (2-seitig) | |
| | | | | Obere | Untere | | | | |
| Paaren 1 | Pcore - Pz3 | ,03600 | ,09008 | ,01645 | ,00236 | ,06964 | 2,189 | 29 | ,037 |
| Paaren 2 | Pcore - Pz2 | ,03533 | ,10207 | ,01864 | -,00278 | ,07345 | 1,896 | 29 | ,068 |
| Paaren 3 | Pz2 - Pz3 | ,00067 | ,08060 | ,01472 | -,02943 | ,03076 | ,045 | 29 | ,964 |
| Paaren 4 | Pcore - Pbaseline | ,02967 | ,06708 | ,01225 | ,00462 | ,05472 | 2,422 | 29 | ,022 |

| Statistik für Test(b) | | | | |
|---|-------------|-------------|-----------|-------------------|
| | Pz3 - Pcore | Pz2 - Pcore | Pz3 - Pz2 | Pbaseline - Pcore |
| Z | -2,035(a) | -1,321(a) | -,336(b) | -2,024(a) |
| Asymptotische Signifikanz (2-seitig) | ,042 | ,187 | ,737 | ,043 |

Signifikanztest KoMoHe IR-Test

Tabelle 81: Statistische Werte und Korrelationen für gemittelte Precision- und Recallwerte (P_CT, P_TT, P_FT, R_CT, R_TT, R_FT) für die 82 Topics der IR-Evaluation KoMoHe. Die Werte wurden paarweise verglichen. Die Auswertungen beziehen sich ausschließlich auf alle Dokumenttypen.

| Statistik bei gepaarten Stichproben | | | | |
|-------------------------------------|------------|----|-------------------------|------------------------------------|
| | Mittelwert | N | Standard- abweichung | Standardfehler des Mittelwertes |
| Paaren 1 | P_CT | 82 | ,28140 | ,03108 |
| | P_TT | 82 | ,26654 | ,02943 |
| Paaren 2 | P_CT | 82 | ,28140 | ,03108 |
| | P_FT | 82 | ,29361 | ,03242 |
| Paaren 3 | P_TT | 82 | ,26654 | ,02943 |
| | P_FT | 82 | ,29361 | ,03242 |
| Paaren 4 | R_CT | 82 | ,38337 | ,04234 |
| | R_TT | 82 | ,33425 | ,03691 |
| Paaren 5 | R_CT | 82 | ,38337 | ,04234 |
| | R_FT | 82 | ,36103 | ,03987 |
| Paaren 6 | R_TT | 82 | ,33425 | ,03691 |
| | R_FT | 82 | ,36103 | ,03987 |

| Korrelationen bei gepaarten Stichproben | | | |
|---|----|-------------|-------------|
| | N | Korrelation | Signifikanz |
| Paaren 1 | 82 | ,387 | ,000 |
| Paaren 2 | 82 | ,273 | ,013 |
| Paaren 3 | 82 | ,207 | ,062 |

| | | | | |
|----------|-------------|----|-------|------|
| Paaren 4 | R_CT & R_TT | 82 | ,450 | ,000 |
| Paaren 5 | R_CT & R_FT | 82 | -,081 | ,467 |
| Paaren 6 | R_TT & R_FT | 82 | -,536 | ,000 |

Tabelle 82: T-Test und Wilcoxon-Test für gemittelte Precisionwerte (P_CT, P_TT, P_FT, R_CT, R_TT, R_FT) für die 82 Topics der IR-Evaluation KoMoHe. Die Signifikanz wurde paarweise berechnet. Die Auswertungen beziehen sich ausschließlich auf alle Dokumenttypen. (a) Basiert auf positiven Rängen. b) Basiert auf negativen Rängen. c) Wilcoxon-Test)

Test bei gepaarten Stichproben

| Gepaarte Differenzen | | | | | | |
|----------------------|-------------|-------------------------|------------------------------------|---|---------|-----------------|
| | | Standard- abweichung | Standardfehler des Mittelwertes | | T | Sig. (2-seitig) |
| | | | Mittelwert | 95% Konfidenzintervall der Differenz | | |
| | | | | Obere | Untere | |
| Paaren 1 | P_CT - P_TT | ,30362 | -,12358 | -,19029 | -,05687 | 81,000 |
| Paaren 2 | P_CT - P_FT | ,34675 | -,26525 | -,34144 | -,18907 | 81,000 |
| Paaren 3 | P_TT - P_FT | ,35333 | -,14168 | -,21931 | -,06404 | 81,000 |
| Paaren 4 | R_CT - R_TT | ,37851 | -,26287 | -,34604 | -,17970 | 81,000 |
| Paaren 5 | R_CT - R_FT | ,54759 | -,23887 | -,35919 | -,11855 | 81,000 |
| Paaren 6 | R_TT - R_FT | ,60937 | ,02400 | -,10990 | ,15789 | 81,722 |

Statistik für Test(b)

| | P_TT - P_CT | P_FT - P_CT | P_FT - P_TT | R_TT - R_CT | R_FT - R_CT | R_FT - R_TT |
|---|-------------|-------------|-------------|-------------|-------------|-------------|
| Z | -3,891(a) | -5,596(a) | -3,044(a) | -5,362(a) | -3,790(a) | -,310(b) |
| Asymptotische Signifikanz (2-seitig) | ,000 | ,000 | ,002 | ,000 | ,000 | ,756 |

Kontrollierte Vokabulare

Tabelle 83: Kontrollierte Vokabulare im Projekt KoMoHe

| Vocabulary | Name | Subject | Type | Mapped | terms | language | CK Database |
|------------|---------|---|------------------|----------|-------|----------|---------------|
| 1 | TheSoz | Thesaurus Sozialwissenschaften | Thesaurus | complete | 7700 | de | 16 SOLIS |
| 2 | Bildung | Thesaurus Bildung | Thesaurus | part | 3500 | de | 2 FIS Bildung |
| 3 | Psyndex | Psyndex Terms | Thesaurus | complete | 5400 | de | 5 Psyndex |
| 4 | DZI | Thesaurus of the Deutsches Institut für soziale Fragen | Thesaurus | complete | 1900 | de | 1 SOLIT |
| 5 | CSA-SA | Thesaurus of Sociological Indexing Terms | Thesaurus | complete | 4300 | en | 1 CSA-SA/SSA |
| 6 | MeSH | Medical Subject Headings | Subject Headings | part | 7500 | en | 4 Medline |
| 7 | BISp | Descriptors of the Bundesinstitut für Sportwissenschaft | Descriptor list | complete | 7400 | de | 4 Spolit |
| 8 | STW | Standard Thesaurus Wirtschaft | Thesaurus | complete | 5700 | de | 3 Econis |
| 9 | IBLK | Thesaurus Internationale Beziehungen und Länderkunde (Euro-Thesaurus) | Thesaurus | complete | 8400 | de | 4 WAO |
| 10 | FES | Descriptors of the Friedrich-Ebert-Stiftung | Descriptor list | complete | 4000 | de | 1 FES DL |
| 11 | GEROLIT | Thesaurus of the Deutsches Zentrum für Altersfragen | Thesaurus | complete | 1900 | de | 2 GEROLIT |
| 12 | SWD | Schlagwortnormdatei | Subject Headings | part | 8800 | de | 2 DNB |

| | | | | | | | | | |
|----|-----------|--|----------------------|-----------------|----------|-------|----|---|------------|
| 13 | CSA-WPSA | CSA Thesaurus of Political Science Indexing Terms | social sciences | Thesaurus | complete | 3100 | en | 1 | WPSA |
| 14 | CSA-PEI | CSA Thesaurus Physical Education Index | sports science | Thesaurus | complete | 1800 | en | 1 | PEI |
| 15 | CSA-ASSIA | CSA Thesaurus Applied Social Sciences Index and Abstracts | social sciences | Thesaurus | complete | 17500 | en | 1 | ASSIA |
| 16 | TWSE | Thesaurus für wirtschaftliche und soziale Entwicklung | political science | Thesaurus | complete | 2800 | de | 1 | INWENT |
| 17 | ELSST | European Language Social Science Thesaurus | social sciences | Thesaurus | complete | 3200 | en | 1 | MADIERA |
| 18 | INFODATA | INFODATA Thesaurus | information science | Thesaurus | complete | 1000 | de | 2 | INFODATA |
| 19 | IAB | Descriptors of the Institut für Arbeitsmarkt- und Berufsforschung | social sciences | Descriptor list | complete | 6800 | de | 2 | IAB |
| 20 | AGROVOC | AGROVOC Thesaurus | agricultural science | Thesaurus | part | 6200 | de | 1 | proArbeit |
| 21 | JEL | Journal of Economic Literature Classification System | economics | Classification | complete | | en | 1 | RePEc |
| 22 | RVK | Regensburger Verbundklassifikation | universal | Classification | part | | de | 1 | UB Regensb |
| 23 | DDC | Dewey Decimal Classification | universal | Classification | part | | de | 1 | WorldCat |
| 24 | CSA-PAIS | CSA Thesaurus PAIS International Subject Headings | political science | Thesaurus | complete | 7400 | en | 1 | PAIS |
| 25 | INION | Descriptors of the Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences | social sciences | Descriptor list | part | | ru | 1 | INION |

Ausschnitte aus der Crosskonkordanz Thesaurus Sozialwissenschaften (TheSoz) zu Standard Thesaurus Wirtschaft (STW)

| TheSoz | Relation | Relevanz | STW |
|---------------------------------------|----------|----------|--|
| Bundeskanzler | = | h | Regierungschef |
| Bundeskompetenz | < | g | Föderalismus |
| Bundeskriminalamt | 0 | | |
| Bundesland | = | h | Teilstaat |
| Bundesministerium | = | h | Ministerium |
| Bundesnachrichtendienst | =+ | h | Staatsschutz + Deutschland |
| Bundespolitik | < | g | Politik |
| Bundespräsident | = | h | Staatsoberhaupt |
| Bundesrat | ^ | m | Parlament |
| Bundesrecht | < | m | Recht |
| Bundesregierung | = | h | Regierung |
| Bundesrepublik Deutschland | = | h | Deutschland |
| Bundesrepublik Jugoslawien | = | h | Serbien-Montenegro |
| Bundessozialgericht | ^ | h | Sozialgericht |
| Bundessozialhilfegesetz | = | h | Sozialhilferecht |
| Bundesstaat | = | m | Föderalismus |
| Bundestag | = | h | Parlament |
| Bundestagswahl | = | h | Wahl |
| Bundesverfassungsgericht | ^ | h | Verfassungsgericht |
| Bundesversammlung | 0 | | |
| Bundesverwaltung | < | g | Öffentliche Verwaltung |
| Bundesverwaltungsgericht | ^ | h | Verwaltungsgericht |
| Bundeswehr | = | h | Militär |
| Bundeszentrale für politische Bildung | 0 | | |
| Bund-Länder-Beziehung | ^ | h | Föderalismus |
| Bund-Länder-Kommission | ^ | g | Föderalismus |
| Bündnis 90/ Die Grünen | =+ | h | Ökologische Partei + Deutschland |
| | | | |
| Eurokommunismus | = | m | Kommunismus |
| Europa | = | h | Europa |
| Europäische Zentralbank | =+ | h | Zentralbank + EU-Staaten |
| Europäer | 0 | | |
| europäische Institution | ^+ | m | Internationale Organisation + EU-Staaten |
| europäische Integration | = | h | Europäische Integration |
| ... | | | |

Ausschnitte aus der Crosskonkordanz Psynindex Terms zu Medical Subject Headings (MeSH)

| Psynindex deutsch | Relation | Relevanz | MeSH |
|----------------------------|----------|----------|---|
| AIDS | =0 | h | Acquired Immunodeficiency Syndrome |
| AIDS | >0 | h | AIDS-Related Complex |
| AIDS | >0 | h | AIDS Dementia Complex |
| AIDS | >0 | m | HIV Wasting Syndrome |
| AIDS | ^0 | h | HIV Infections |
| AIDS-Demenz | = | h | AIDS Dementia Complex |
| AIDS-Prävention | < | m | Acquired Immunodeficiency Syndrome |
| Akademisches Fachpersonal | <0 | m | Occupational Groups |
| Akademisches Fachpersonal | <0 | m | Faculty |
| Akademisches Fachpersonal | ^0 | g | Professional Role |
| Akademisches Fachpersonal | ^0 | m | Education, Professional |
| Akathisie | ^ | h | Psychomotor Agitation |
| Akkulturation | = | h | Acculturation |
| Akrophobie | < | m | Phobic Disorders |
| Aktivismus | < | m | Politics |
| Aktivitätsniveau | < | g | Human Activities |
| Aktivitätstheorie | <+ | g | Human Activities + Aging |
| Akupunktur | = | h | Acupuncture |
| Akustik | = | h | Acoustics |
| Akustische Displays | < | g | Acoustic Stimulation |
| Akustische Halluzinationen | =+ | h | Acoustics + Hallucinations |
| Akustischer Nerv | = | h | Cochlear Nerve |
| Akustischer Reflex | = | h | Reflex, Acoustic |
| Akute Alkoholvergiftung | = | h | Ethanol/poisoning |
| Akute Psychose | < | h | Psychotic Disorders |
| Akute Schizophrenie | =0 | h | Schizophrenia |
| Akute Schizophrenie | ^0 | m | Schizophrenia and Disorders with Psychotic Features |
| Akute Schizophrenie | ^0 | m | Schizophrenia, Disorganized |
| Akute Schizophrenie | ^0 | m | Schizophrenia, Catatonic |
| Akute Schizophrenie | ^0 | m | Schizophrenia, Childhood |
| Akute Schizophrenie | ^0 | m | Schizophrenia, Paranoid |
| Akute Stressstörung | ^ | h | Stress Disorders, Traumatic, Acute |
| Alanine | = | h | Alanine |
| Alarmreaktionen | ^ | m | Escape Reaction |
| Alaska-Bevölkerung | < | m | Alaska |
| ... | | | |

Indirekte Term-Transformationen

Tabelle 84: ÄQ-Relationen, TheSoz > Gerolit (1.TT). Gerolit > MESH (2.TT), ausgehend vom Switching Term

| Startterm (TheSoz) | Switching Term (Gerolit) | Endterm (MESH) |
|-------------------------|----------------------------|----------------------------|
| Allergie | Allergische Erkrankungen | Hypersensitivity |
| älterer Arbeitnehmer | Ältere Erwerbstätige | Middle Aged + Employment |
| Alterssoziologie | Soziologische Gerontologie | Sociology + Geriatrics |
| Anomie | Abweichendes Verhalten | Behavioral Symptoms |
| Arbeitnehmer | Erwerbstätige | Employment |
| Arbeitspsychologie | Arbeitswissenschaft | Human Engineering |
| Ausbildungsstand | Bildungsstand | Educational Status |
| Autonomie | Selbständigkeit | Personal Autonomy |
| Behindertenwerkstätte | Beschützende Werkstätte | Sheltered Workshops |
| Beratungsgremium | Beirat | Advisory Committees |
| Berufsmobilität | Berufliche Mobilität | Career Mobility |
| Berufstätigkeit | Erwerbstätigkeit | Employment |
| Beschäftigungstherapeut | Ergotherapeuten | Occupational Therapy/MA |
| Bildungsniveau | Bildungsstand | Educational Status |
| Bildungsprogramm | Bildungsplanung | Education + Public Policy |
| biographische Methode | Biographische Analyse | Biography |
| Eigenarbeit | Schattenwirtschaft | Economics |
| Einsparung | Sparmaßnahmen | Cost Savings |
| Einwanderung | Migration | Emigration and Immigration |
| Emotionalität | Emotionen | Emotions |

Tabelle 85: ÄQ-Relationen, STW > IBLK (1.TT). IBLK > CSA-PAIS (2.TT), ausgehend vom Switching Term

| Startterm (STW) | Switching Term (IBLK) | Endterm (CSA-PAIS) |
|-----------------------------|---|----------------------------------|
| Altenhilfe | Altenbetreuung/Altenhilfe | Elder care |
| Alterssicherung | Altersversorgung | Old age pensions |
| Alterssicherung | Altersversorgung | Retirement income |
| Anarchie | Anarchismus | Anarchism |
| Anbausystem | Fruchtwechsel | Crop rotation |
| Antizyklische Finanzpolitik | Öffentliche Finanzpolitik | Public finance + Economic policy |
| Arbeitsmarktforschung | Arbeitsmarktforschung/ Berufsforschung | Labor + Research |
| Arbeitszeitflexibilisierung | Flexible Arbeitszeit | Flexible schedules |
| Architekturbüro | Architektur | Architecture |
| Asylrecht | Recht auf Asyl | Asylum, Right of |
| Auslandsvermögen | Auslandseigentum | Alien property |
| Auslandsverschuldung | Externe Verschuldung | Debts, External |
| Auslandsverschuldung | Externe Verschuldung | External debts |
| Basisgesundheitsversorgung | Medizinische Versorgung | Medical care |
| Beamte | Öffentlich Bedienstete | Public officials |
| Befristeter Arbeitsvertrag | Befristete Beschäftigung | Temporary employment |
| Bildungsabschluss | Bildungsabschluss/ Abschlussprüfung | Educational attainment |
| Biodiversität | Biologische Vielfalt | Biodiversity |
| Branche | Wirtschaftszweige | Branches |
| Bündnissystem | Verteidigungsbündnis | Defenses + Alliances |

Tabelle 86: ÄQ-Relationen, TheSoz > Psyndex (1.TT). Psyndex > MESH (2.TT), ausgehend vom Switching Term

| Startterm (TheSoz) | Switching Term (Psyndex) | Endterm (MESH) |
|---------------------------------|---------------------------------------|--------------------------------|
| Algorithmus | Algorithmen | Algorithms |
| Berufsethos | Berufsethik | Ethics, Business |
| Berufsmobilität | Berufliche Mobilität | Career Mobility |
| Cluster-Analyse | Clusteranalyse | Cluster Analysis |
| Datengewinnung | Datensammlung | Data Collection |
| experimentelle Psychologie | Experimentalpsychologie | Psychology, Experimental |
| gesetzliche Krankenversicherung | Krankenversicherung von Arbeitnehmern | Health Benefit Plans, Employee |
| Hilfeleistung | Hilfeverhalten | Helping Behavior |
| Hochbegabung | Hochbegabte | Child, Gifted |
| Ich-Störung | Depersonalisation | Depersonalization |
| Impfung | Immunisierung | Immunization |
| Justizvollzugsanstalt | Gefängnisse | Prisons |
| Kinderlähmung | Poliomyelitis | Poliomyelitis |
| Knochenkrankheit | Osteoporose | Osteoporosis |
| Lernschwierigkeit | Lernstörungen | Learning Disorders |
| Rechtsanwalt | Anwälte | Lawyers |

Tabelle 87: ÄQ-Relationen, TheSoz > IBLK (1.TT). IBLK > CSA-PAIS (2.TT), ausgehend vom Switching Term

| Startterm (TheSoz) | Switching Term (IBLK) | Endterm (CSA-PAIS) |
|---------------------|--|---------------------------|
| Agrarbetrieb | Landwirtschaftlicher Betrieb | Farms |
| Angebotspolitik | Angebotsorientierte Wirtschaftspolitik | Supply-side economics |
| Angebotstheorie | Angebotsorientierte Wirtschaftstheorie | Supply-side economics |
| anglophones Afrika | Englischsprachiges Afrika | English-speaking Africa |
| Arabische Liga | League of Arab States | Arab league |
| Arbeiterschaft | Arbeiter | Employees |
| Arbeiterschaft | Arbeiter | Labor |
| Aussiedlung | Umsiedlung | Resettlement |
| Beamter | Öffentlich Bedienstete | Public officials |
| Biodiversität | Biologische Vielfalt | Biodiversity |
| Corporate Identity | Unternehmenskultur | Corporate culture |
| Demographie | Bevölkerungswissenschaft | Demographics |
| Demographie | Bevölkerungswissenschaft | Population |
| Demonstration | Öffentliche Demonstration | Demonstrations, Political |
| E-Mail | Electronic Mail | Electronic mail |
| Empfängnisverhütung | Schwangerschaftsverhütung | Pregnancy + Prevention |
| Erfindung | Invention | Inventions |
| Facharbeiter | Facharbeitskräfte | Skilled labor |
| Familienangehöriger | Familie | Family |
| Familieneinkommen | Familienbudget | Budget, Family |

Tabelle 88: ÄQ-Relationen, CSA-SA > TheSoz (1.TT). TheSoz > STW (2.TT), ausgehend vom Switching Term

| Startterm (CSA-SA) | Switching Term (TheSoz) | Endterm (STW) |
|---------------------------|--------------------------------|--------------------------------|
| Abortion | Schwangerschaftsabbruch | Schwangerschaftsabbruch |
| Labor Force Participation | Erwerbsbeteiligung | Erwerbstätigkeit |
| Labor Theory of Value | Arbeitswerttheorie | Arbeitswertlehre |
| Management Styles | Führungsstil | Führungstheorie |
| Manufacturing Industries | produzierendes Gewerbe | Industrie |
| Mass Media | Massenmedien | Kommunikationsmedien |
| Measures (Instruments) | Messinstrument | Messgerät |
| Metropolitan Areas | Ballungsgebiet | Ballungsraum |
| Middle Class | Mittelschicht | Mittelstand |
| Needs | Bedürfnis | Bedürfnisse |
| Nuclear Weapons | Kernwaffe | Atomwaffe |
| Parks | Grünfläche | Städtische Grünfläche |
| Part Time Farming | Nebenerwerbsbetrieb | Nebenerwerbslandwirtschaft |
| Professional Associations | Berufsverband | Berufsverband |
| Religious Orders | Orden | Religiöser Orden |
| Social Consciousness | gesellschaftliches Bewusstsein | gesellschaftliches Bewusstsein |

In dieser Arbeit werden zwei Mehrwertdienste für Suchsysteme vorgestellt, die typische Probleme bei der Recherche nach wissenschaftlicher Literatur behandeln können. Die beiden Mehrwertdienste semantische Heterogenitätsbehandlung am Beispiel Crosskonkordanzen und Re-Ranking auf Basis von Bradfordizing, die in unterschiedlichen Phasen der Suche zum Einsatz kommen, werden in diesem Buch ausführlich beschrieben und evaluiert.

Für die Tests wurden Fragestellungen und Daten aus zwei Evaluationsprojekten (CLEF und KoMoHe) verwendet. Die intellektuell bewerteten Dokumente stammen aus insgesamt sieben Fachdatenbanken der Fächer Sozialwissenschaften, Politikwissenschaft, Wirtschaftswissenschaften, Psychologie und Medizin. Die Ergebnisse dieser Arbeit sind in das GESIS-Projekt IRM eingeflossen.

This thesis proposes two value-added approaches for search systems which treat typical problems in searching scientific literature and seek to improve the retrieval situation on a measurable level. The two value-added services, semantic treatment of heterogeneity (the example of cross-concordances) and re-ranking on Bradfordizing, which are applied in different search phases, are described in detail and their effectiveness in typical subject-specific searches is evaluated in the empirical part of the thesis.

We used topics and data from two evaluation projects (CLEF and KoMoHe) for the tests. The intellectually assessed documents come from seven academic abstracting and indexing databases representing social science, political science, economics, psychology and medicine. The results of this work will be integrated in the GESIS project IRM.